



## EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention  
of the grant of the patent:  
30.08.2006 Bulletin 2006/35

(21) Application number: 01922524.2

(22) Date of filing: 21.03.2001

(51) Int Cl.:

*C12N 15/63* (2006.01) *C12N 15/65* (2006.01)  
*C12N 15/12* (2006.01) *C12N 15/62* (2006.01)  
*C07K 14/705* (2006.01) *C07K 14/72* (2006.01)  
*A01H 5/00* (2006.01) *C12Q 1/68* (2006.01)  
*G01N 33/68* (2006.01)

(86) International application number:  
PCT/US2001/009050

(87) International publication number:  
WO 2001/070816 (27.09.2001 Gazette 2001/39)

(54) **ECDYSONE RECEPTOR-BASED INDUCIBLE GENE EXPRESSION SYSTEM**

INDUZIERTES GENEXPRESSIONSSYSTEM BASIEREND AUF ECDYSONREZEPTOREN  
NOUVEAU SYSTEME D'EXPRESSION GENETIQUE INDUCIBLE FONDE SUR LE LE RECEPTEUR  
A L'ECDYSONE

(84) Designated Contracting States:  
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU  
MC NL PT SE TR

(30) Priority: 22.03.2000 US 191355 P  
20.02.2001 US 269799 P

(43) Date of publication of application:  
18.12.2002 Bulletin 2002/51

(73) Proprietor: RheoGene Holdings Inc.  
Norristown, PA 19403 (US)

(72) Inventors:  
• PALLI, Subba, Reddy  
Lansdale, PA 19446 (US)  
• KAPITSKAYA, Marianna Zinovjevna  
North Wales, PA 19454 (US)  
• CRESS, Dean, Ervin  
Sounderton, PA 18964 (US)

(74) Representative: Harding, Charles Thomas et al  
D Young & Co  
120 Holborn  
London EC1N 2DY (GB)

(56) References cited:  
EP-A- 0 798 378 EP-A- 0 965 644  
WO-A-96/27673 WO-A-99/02683  
WO-A-99/10510 WO-A-99/58155

- CHRISTOPHERSON K S ET AL: "ECDYSTEROID-DEPENDENT REGULATION OF GENES IN MAMMALIAN CELLS BY A DROSOPHILA ECDYSONE RECEPTOR AND CHIMERIC TRANSACTIVATORS" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE. WASHINGTON, US, vol. 89, July 1992 (1992-07), pages 6314-6318, XP002921831 ISSN: 0027-8424
- KAKIZAWA TOMOKO ET AL: "Ligand-dependent heterodimerization of thyroid hormone receptor and retinoid X receptor." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 272, no. 38, 1997, pages 23799-23804, XP002200320 ISSN: 0021-9258
- PERERA SRINI C ET AL: "An analysis of ecdysone receptor domains required for heterodimerization with ultraspiracle." ARCHIVES OF INSECT BIOCHEMISTRY AND PHYSIOLOGY, vol. 41, no. 2, 1999, pages 61-70, XP001077866 ISSN: 0739-4462
- KOELLE M R ET AL: "THE DROSOPHILA ECR GENE ENCODES AN ECDYSONE RECEPTOR, A NEW MEMBER OF THE STEROID RECEPTOR SUPERFAMILY" CELL, CELL PRESS, CAMBRIDGE, MA, US, vol. 67, no. 1, 4 October 1991 (1991-10-04), pages 59-77, XP001010154 ISSN: 0092-8674

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

- LEID M ET AL: "PURIFICATION, CLONING, AND RXR IDENTITY OF THE HELA CELL FACTOR WITH WHICH RAR OR TR HETERODIMERIZES TO BIND TARGET SEQUENCES EFFICIENTLY" CELL, CELL PRESS, CAMBRIDGE, NA, US, vol. 68, 24 January 1992 (1992-01-24), pages 377-395, XP002007203 ISSN: 0092-8674
- PERERA SRINI C ET AL: "Studies on two ecdysone receptor isoforms of the spruce budworm, *Choristoneura fumiferana*." MOLECULAR AND CELLULAR ENDOCRINOLOGY, vol. 152, no. 1-2, 25 June 1999 (1999-06-25), pages 73-84, XP002200321 ISSN: 0303-7207

**Description****FIELD OF THE INVENTION**

[0001] This invention relates to the field of biotechnology or genetic engineering. Specifically, this invention relates to the field of gene expression. More specifically, this invention relates to a novel ecdysone receptor-based inducible gene expression system and methods of modulating the expression of a gene within a host cell using this inducible gene expression system.

**BACKGROUND OF THE INVENTION**

[0002] In the field of genetic engineering, precise control of gene expression is a valuable tool for studying, manipulating, and controlling development and other physiological processes. Gene expression is a complex biological process involving a number of specific protein-protein interactions. In order for gene expression to be triggered, such that it produces the RNA necessary as the first step in protein synthesis, a transcriptional activator must be brought into proximity of a promoter that controls gene transcription. Typically, the transcriptional activator itself is associated with a protein that has at least one DNA binding domain that binds to DNA binding sites present in the promoter regions of genes. Thus, for gene expression to occur, a protein comprising a DNA binding domain and a transactivation domain located at an appropriate distance from the DNA binding domain must be brought into the correct position in the promoter region of the gene.

[0003] The traditional transgenic approach utilizes a cell-type specific promoter to drive the expression of a designed transgene. A DNA construct containing the transgene is first incorporated into a host genome. When triggered by a transcriptional activator, expression of the transgene occurs in a given cell type.

[0004] Another means to regulate expression of foreign genes in cells is through inducible promoters. Examples of the use of such inducible promoters include the **PR1-a** promoter, prokaryotic repressor-operator systems, immunosuppressive-immunophilin systems, and higher eukaryotic transcription activation systems such as steroid hormone receptor systems and are described below.

[0005] The **PR1-a** promoter from tobacco is induced during the systemic acquired resistance response following pathogen attack. The use of **PR1-a** may be limited because it often responds to endogenous materials and external factors such as pathogens, UV-B radiation, and pollutants. Gene regulation systems based on promoters induced by heat shock, interferon and heavy metals have been described (Wurn et al., 1986, *Proc. Natl. Acad. Sci. USA* 83: 5414-5418; Arnheiter et al., 1990 *Cell* 62:51-61; Filmus et al., 1992 *Nucleic Acids Research* 20:27550-27560). However, these systems have limitations due to their effect on expression of non-target genes. These systems are also leaky.

[0006] Prokaryotic repressor-operator systems utilize bacterial repressor proteins and the unique operator DNA sequences to which they bind. Both the tetracycline ("Tet") and lactose ("Lac") repressor-operator systems from the bacterium *Escherichia coli* have been used in plants and animals to control gene expression. In the Tet system, tetracycline binds to the TetR repressor protein, resulting in a conformational change which releases the repressor protein from the operator which as a result allows transcription to occur. In the Lac system, a lac operon is activated in response to the presence of lactose, or synthetic analogs such as isopropyl-b-D-thiogalactoside. Unfortunately, the use of such systems is restricted by unstable chemistry of the ligands, *i.e.* tetracycline and lactose, their toxicity, their natural presence, or the relatively high levels required for induction or repression. For similar reasons, utility of such systems in animals is limited.

[0007] Immunosuppressive molecules such as FK506, rapamycin and cyclosporine A can bind to immunophilins FKBP12, cyclophilin, *etc.* Using this information, a general strategy has been devised to bring together any two proteins simply by placing FK506 on each of the two proteins or by placing FK506 on one and cyclosporine A on another one. A synthetic homodimer of FK506 (FK1012) or a compound resulted from fusion of FK506-cyclosporine (FKCsA) can then be used to induce dimerization of these molecules (Spencer et al., 1993, *Science* 262:1019-24; Belshaw et al., 1996 *Proc Natl Acad Sci U S A* 93:4604-7). Gal4 DNA binding domain fused to FKBP12 and VP16 activator domain fused to cyclophilin, and FKCsA compound were used to show heterodimerization and activation of a reporter gene under the control of a promoter containing Gal4 binding sites. Unfortunately, this system includes immunosuppressants that can have unwanted side effects and therefore, limits its use for various mammalian gene switch applications.

[0008] Higher eukaryotic transcription activation systems such as steroid hormone receptor systems have also been employed. Steroid hormone receptors are members of the nuclear receptor superfamily and are found in vertebrate and invertebrate cells. Unfortunately, use of steroidal compounds that activate the receptors for the regulation of gene expression, particularly in plants and mammals, is limited due to their involvement in many other natural biological pathways in such organisms. In order to overcome such difficulties, an alternative system has been developed using insect ecdysone receptors (EcR).

[0009] Growth, molting, and development in insects are regulated by the ecdysone steroid hormone (molting hormone)

and the juvenile hormones (Dhadialla, et al., 1998, Annu. Rev. Entomol. 43: 545-569). The molecular target for ecdysone in insects consists of at least ecdysone receptor (EcR) and ultraspiracle protein (USP). EcR is a member of the nuclear steroid receptor super family that is characterized by signature DNA and ligand binding domains, and an activation domain (Koelle et al. 1991, Cell, 67:59-77). EcR receptors are responsive to a number of steroidal compounds such as ponasterone A and muristerone A. Recently, non-steroidal compounds with ecdysteroid agonist activity have been described, including the commercially available insecticides tebufenozide and methoxyfenozide that are marketed world wide by Rohm and Haas Company (see International Patent Application No. PCT/EP96/00686 (WO 96/27673) and US Patent 5,530,028). Both analogs have exceptional safety profiles to other organisms.

[0010] International Patent Application No. PCT/US97/05330 (WO 97/38117) discloses methods for modulating the expression of an exogenous gene in which a DNA construct comprising the exogenous gene and an ecdysone response element is activated by a second DNA construct comprising an ecdysone receptor that, in the presence of a ligand therefor, and optionally in the presence of a receptor capable of acting as a silent partner, binds to the ecdysone response element to induce gene expression. The ecdysone receptor of choice was isolated from *Drosophila melanogaster*. Typically, such systems require the presence of the silent partner, preferably retinoid X receptor (RXR), in order to provide optimum activation. In mammalian cells, insect ecdysone receptor (EcR) heterodimerizes with retinoid X receptor (RXR) and regulates expression of target genes in a ligand dependent manner. International Patent Application No. PCT/US98/14215 (WO 99/02683) discloses that the ecdysone receptor isolated from the silk moth *Bombyx mori* is functional in mammalian systems without the need for an exogenous dimer partner.

[0011] U.S. Patent No. 5,880,333 discloses a *Drosophila melanogaster* EcR and ultraspiracle (USP) heterodimer system used in plants in which the transactivation domain and the DNA binding domain are positioned on two different hybrid proteins. Unfortunately, this system is not effective for inducing reporter gene expression in animal cells (for comparison, see Example 1.2, below).

[0012] In each of these cases, the transactivation domain and the DNA binding domain (either as native EcR as in International Patent Application No. PCT/US98/14215 or as modified EcR as in International Patent Application No. PCT/US97/05330) were incorporated into a single molecule and the other heterodimeric partners, either USP or RXR, were used in their full length form.

[0013] Drawbacks of the above described EcR-based gene regulation systems include a considerable background activity in the absence of ligands and that these systems are not applicable for use in both plants and animals (see U.S. Patent No. 5,880,333). For most applications that rely on modulating gene expression, these EcR-based systems are undesirable. Therefore, a need exists in the art for improved systems to precisely modulate the expression of exogenous genes in both plants and animals. Such improved systems would be useful for applications such as gene therapy, large scale production of proteins and antibodies, cell-based high throughput screening assays, functional genomics and regulation of traits in transgenic animals. Improved systems that are simple, compact, and dependent on ligands that are relatively inexpensive, readily available, and of low toxicity to the host would prove useful for regulating biological systems.

[0014] Various publications are cited herein, the disclosures of which are incorporated by reference in their entireties. However, the citation of any reference herein should not be construed as an admission that such reference is available as "Prior Art" to the instant application.

## **SUMMARY OF THE INVENTION**

[0015] The present invention relates to a novel ecdysone receptor-based inducible gene expression system, novel receptor polynucleotides and polypeptides for use in the novel inducible gene expression system, and methods of modulating the expression of a gene within a host cell using this inducible gene expression system. In particular, Applicants' invention relates to an improved gene expression modulation system comprising a polynucleotide encoding a receptor polypeptide comprising a truncation mutation.

[0016] Specifically, the present invention relates to a gene expression modulation system comprising: a) a first gene expression cassette that is capable of being expressed in a host cell comprising a polynucleotide that encodes a first polypeptide comprising: i) a DNA-binding domain that recognizes a response element associated with a gene whose expression is to be modulated; and ii) a ligand binding domain comprising a ligand binding domain from a nuclear receptor; and b) a second gene expression cassette that is capable of being expressed in the host cell comprising a polynucleotide sequence that encodes a second polypeptide comprising: i) a transactivation domain; and ii) a ligand binding domain comprising a ligand binding domain from a nuclear receptor other than an ultraspiracle receptor; wherein the DNA binding domain and the transactivation domain are from a polypeptide other than an ecdysone receptor, a retinoid X receptor, or an ultraspiracle receptor, wherein the ligand binding domains from the first polypeptide and the second polypeptide are different and dimerize.

[0017] In a specific embodiment, the ligand binding domain of the first polypeptide comprises an ecdysone receptor (EcR) ligand binding domain

[0018] In another specific embodiment, the ligand binding domain of the second polypeptide comprises a retinoid X receptor (RXR) ligand binding domain.

[0019] In a preferred embodiment, the ligand binding domain of the first polypeptide comprises an ecdysone receptor ligand binding domain and the ligand binding domain of the second polypeptide comprises a retinoid X receptor ligand binding domain

[0020] The present invention also relates to a gene expression modulation system according to the invention further comprising c) a third gene expression cassette comprising: i) a response element to which the DNA-binding domain of the first polypeptide binds; ii) a promoter that is activated by the transactivation domain of the second polypeptide; and iii) the gene whose expression is to be modulated.

[0021] The present invention also relates to an isolated polynucleotide encoding a truncated EcR or a truncated RXR polypeptide, wherein the truncation mutation affects ligand binding activity or ligand sensitivity.

[0022] In particular, the present invention relates to an isolated polynucleotide encoding a truncated EcR or a truncated RXR polypeptide comprising a truncation mutation that reduces ligand binding activity or ligand sensitivity of said EcR or RXR polypeptide. In a specific embodiment, the present invention relates to an isolated polynucleotide encoding a truncated EcR or a truncated RXR polypeptide comprising a truncation mutation that reduces steroid binding activity or steroid sensitivity of said EcR or RXR polypeptide. In another specific embodiment, the present invention relates to an isolated polynucleotide encoding a truncated EcR or a truncated RXR polypeptide comprising a truncation mutation that reduces non-steroid binding activity or non-steroid sensitivity of said EcR or RXR polypeptide.

[0023] The present invention also relates to an isolated polynucleotide encoding a truncated EcR or a truncated RXR polypeptide comprising a truncation mutation that enhances ligand binding activity or ligand sensitivity of said EcR or RXR polypeptide. In a specific embodiment, the present invention relates to an isolated polynucleotide encoding a truncated EcR or a truncated RXR polypeptide comprising a truncation mutation that enhances steroid binding activity or steroid sensitivity of said EcR or RXR polypeptide. In another specific embodiment, the present invention relates to an isolated polynucleotide encoding a truncated EcR or a truncated RXR polypeptide comprising a truncation mutation that enhances non-steroid binding activity or non-steroid sensitivity of said EcR or RXR polypeptide.

[0024] The present invention also relates to an isolated polynucleotide encoding a truncated RXR polypeptide comprising a truncation mutation that increases ligand sensitivity of a heterodimer comprising the truncated retinoid X receptor polypeptide and a dimerization partner. In a specific embodiment, the dimerization partner is an ecdysone receptor polypeptide.

[0025] The present invention also relates to an isolated polypeptide encoded by a polynucleotide according to Applicants' invention. In particular, the present invention relates to an isolated truncated EcR or truncated RXR polypeptide comprising a truncation mutation, wherein the EcR or RXR polypeptide is encoded by a polynucleotide according to the invention.

[0026] Thus, the present invention also relates to an isolated truncated EcR or truncated RXR polypeptide comprising a truncation mutation that affects ligand binding activity or ligand sensitivity of said EcR or RXR polypeptide.

[0027] Applicants' invention also relates to methods of modulating gene expression in a host cell using a gene expression modulation system according to the invention. Specifically, Applicants' invention provides a method of modulating the expression of a gene in a host cell comprising the gene to be modulated comprising the steps of: a) introducing into the host cell a gene expression modulation system according to the invention; and b) introducing into the host cell a ligand that independently combines with the ligand binding domains of the first polypeptide and the second polypeptide of the gene expression modulation system; wherein the gene to be expressed is a component of a chimeric gene comprising: i) a response element comprising a domain to which the DNA binding domain from the first polypeptide binds; ii) a promoter that is activated by the transactivation domain of the second polypeptide; and iii) the gene whose expression is to be modulated, whereby a complex is formed comprising the ligand, the first polypeptide, and the second polypeptide, and whereby the complex modulates expression of the gene in the host cell.

[0028] Applicants' invention also provides an isolated host cell comprising an inducible gene expression system according to the invention. The present invention also relates to an isolated host cell comprising a polynucleotide or polypeptide according to the invention. Accordingly, Applicants' invention also relates to a non-human organism comprising a host cell according to the invention.

## BRIEF DESCRIPTION OF THE DRAWINGS

[0029]

Figure 1: An ecdysone receptor-based gene expression system comprising a first gene expression cassette encoding a Gal4DBD-CfEcRDEF chimeric polypeptide and a second gene expression cassette encoding a VP16AD-MmRXR-DEF chimeric polypeptide; prepared as described in Example 1 (switch 1.1).

Figure 2: An ecdysone receptor-based gene expression system comprising a first gene expression cassette encoding

a Gal4DBD-CfEcrDEF chimeric polypeptide and a second gene expression cassette encoding a VP16AD-CfUSP-DEF chimeric polypeptide; prepared as described in Example 1 (switch 1.2).

Figure 3: An ecdysone receptor-based gene expression system comprising a first gene expression cassette encoding a Gal14DBD-MmRXRDEF chimeric polypeptide and a second gene expression cassette encoding a VP16AD-CfEcrDEF chimeric polypeptide; prepared as described in Example 1 (switch 1.3).

Figure 4: An ecdysone receptor-based gene expression system comprising a first gene expression cassette encoding a Gal14DBD-MmRXRDEF chimeric polypeptide and a second gene expression cassette encoding a VP16AD-DmEcrDEF chimeric polypeptide; prepared as described in Example 1 (switch 1.4).

Figure 5: An ecdysone receptor-based gene expression system comprising a first gene expression cassette encoding a Gal14DBD-CfUSPDEF chimeric polypeptide and a second gene expression cassette encoding a VP16AD-CfEcrDEF chimeric polypeptide; prepared as described in Example 1 (switch 1.5).

Figure 6: An ecdysone receptor-based gene expression system comprising a first gene expression cassette encoding a Gal14DBD-CfEcrDEF-VP16AD chimeric polypeptide; prepared as described in Example 1 (switch 1.6).

Figure 7: An ecdysone receptor-based gene expression system comprising a first gene expression cassette encoding a VP16AD-CfEcrDEF chimeric polypeptide; prepared as described in Example 1 (switch 1.7).

Figure 8: An ecdysone receptor-based gene expression system comprising a first gene expression cassette encoding a VP16AD-DmEcrDEF chimeric polypeptide and a second gene expression cassette encoding a MmRXR polypeptide; prepared as described in Example 1 (switch 1.8).

Figure 9: An ecdysone receptor-based gene expression system comprising a first gene expression cassette encoding a VP16AD-CfEcrDEF chimeric polypeptide and a second gene expression cassette encoding a MmRXR polypeptide; prepared as described in Example 1 (switch 1.9).

Figure 10: An ecdysone receptor-based gene expression system comprising a gene expression cassette encoding a Gal4DBD-CfEcrDEF chimeric polypeptide; prepared as described in Example 1 (switch 1.10).

Figure 11: Expression data of GAL4CfEcrA/BCDEF, GAL4CfEcrCDEF, GAL4CfEcr1/2CDEF, GAL4CfEcrDEF, GAL4CfEcrEF, GAL4CfEcrDE truncation mutants transfected into NIH3T3 cells along with VP16MmRXRDE, pFRLuc and pTKRL plasmid DNAs.

Figure 12: Expression data of GAL4CfEcrA/BCDEF, GAL4CfEcrCDEF, GAL4CfEcr1/2CDEF, GAL4CfEcrDEF, GAL4CfEcrEF, GAL4CfEcrDE truncation mutants transfected into 3T3 cells along with VP16MmRXRE, pFRLuc and pTKRL plasmid DNAs.

Figure 13: Expression data of VP16MmRXRA/BCDEF, VP16MmRXRCDEF, VP16MmRXRDEF, VP16MmRXREF, VP16MmRXRBam-EF, VP16MmRXRAF2del constructs transfected into NIH3T3 cells along with GAL4CfEcrDEF, pFRLuc and pTKRL plasmid DNAs.

Figure 14: Expression data of VP16MmRXRA/BCDEF, VP16MmRXRCDEF, VP16MmRXRDEF, VP16MmRXREF, VP16MmRXRBam-EF, VP16MmRXRAF2del constructs transfected into NIH3T3 cells along with GAL4CfEcrDEF, pFRLuc and pTKRL plasmid DNAs.

Figure 15: Expression data of various truncated CfEcr and MmRXR receptor pairs transfected into NIH3T3 cells along with GAL4CfEcrDEF, pFRLuc and pTKRL plasmid DNAs.

## DETAILED DESCRIPTION OF THE INVENTION

[0030] Applicants have now developed an improved ecdysone receptor-based inducible gene expression system comprising a truncation mutant of an ecdysone receptor or a retinoid X receptor (RXR) polypeptide that affects ligand binding activity or ligand sensitivity. This mutational effect may increase or reduce ligand binding activity or ligand sensitivity and may be steroid or non-steroid specific. Thus, Applicants' invention provides an improved ecdysone receptor-based inducible gene expression system useful for modulating expression of a gene of interest in a host cell. In a particularly desirable embodiment, Applicants' invention provides an inducible gene expression system that has a reduced level of background gene expression and responds to submicromolar concentrations of non-steroidal ligand. Thus, Applicants' novel inducible gene expression system and its use in methods of modulating gene expression in a host cell overcome the limitations of currently available inducible expression systems and provide the skilled artisan with an effective means to control gene expression.

[0031] The present invention provides a novel inducible gene expression system that can be used to modulate gene expression in both prokaryotic and eukaryotic host cells. Applicants' invention is useful for applications such as gene therapy, large scale production of proteins and antibodies, cell-based high throughput screening assays, functional genomics and regulation of traits in transgenic organisms.

## DEFINITIONS

[0032] In this disclosure, a number of terms and abbreviations are used. The following definitions are provided and

should be helpful in understanding the scope and practice of the present invention.

**[0033]** In a specific embodiment, the term "about" or "approximately" means within 20%, preferably within 10%, more preferably within 5%, and even more preferably within 1% of a given value or range.

**[0034]** The term "substantially free" means that a composition comprising "A" (where "A" is a single protein, DNA molecule, vector, recombinant host cell, etc.) is substantially free of "B" (where "B" comprises one or more contaminating proteins, DNA molecules, vectors, etc.) when at least about 75% by weight of the proteins, DNA, vectors (depending on the category of species to which A and B belong) in the composition is "A". Preferably, "A" comprises at least about 90% by weight of the A+B species in the composition, most preferably at least about 99% by weight. It is also preferred that a composition, which is substantially free of contamination, contain only a single molecular weight species having the activity or characteristic of the species of interest.

**[0035]** The term "isolated" for the purposes of the present invention designates a biological material (nucleic acid or protein) that has been removed from its original environment (the environment in which it is naturally present).

**[0036]** For example, a polynucleotide present in the natural state in a plant or an animal is not isolated. The same polynucleotide separated from the adjacent nucleic acids in which it is naturally present. The term "purified" does not require the material to be present in a form exhibiting absolute purity, exclusive of the presence of other compounds. It is rather a relative definition.

**[0037]** A polynucleotide is in the "purified" state after purification of the starting material or of the natural material by at least one order of magnitude, preferably 2 or 3 and preferably 4 or 5 orders of magnitude.

**[0038]** A "nucleic acid" is a polymeric compound comprised of covalently linked subunits called nucleotides. Nucleic acid includes polyribonucleic acid (RNA) and polydeoxyribonucleic acid (DNA), both of which may be single-stranded or double-stranded. DNA includes but is not limited to cDNA, genomic DNA, plasmids DNA, synthetic DNA, and semi-synthetic DNA. DNA may be linear, circular, or supercoiled.

**[0039]** A "nucleic acid molecule" refers to the phosphate ester polymeric form of nucleosides (adenosine, guanosine, uridine or cytosine; "RNA molecules") or deoxyribonucleosides (deoxyadenosine, deoxyguanosine, deoxythymidine, or deoxycytidine; "DNA molecules"), or any phosphoester analogs thereof, such as phosphorothioates and thioesters, in either single stranded form, or a double-stranded helix. Double stranded DNA-DNA, DNA-RNA and RNA-RNA helices are possible. The term nucleic acid molecule, and in particular DNA or RNA molecule, refers only to the primary and secondary structure of the molecule, and does not limit it to any particular tertiary forms. Thus, this term includes double-stranded DNA found, *inter alia*, in linear or circular DNA molecules (e.g., restriction fragments), plasmids, and chromosomes. In discussing the structure of particular double-stranded DNA molecules, sequences may be described herein according to the normal convention of giving only the sequence in the 5' to 3' direction along the non-transcribed strand of DNA (*i.e.*, the strand having a sequence homologous to the mRNA). A "recombinant DNA molecule" is a DNA molecule that has undergone a molecular biological manipulation.

**[0040]** The term "fragment" will be understood to mean a nucleotide sequence of reduced length relative to the reference nucleic acid and comprising, over the common portion, a nucleotide sequence identical to the reference nucleic acid. Such a nucleic acid fragment according to the invention may be, where appropriate, included in a larger polynucleotide of which it is a constituent. Such fragments comprise, or alternatively consist of, oligonucleotides ranging in length from at least 8, 10, 12, 15, 18, 20 to 25, 30, 40, 50, 70, 80, 100, 200, 500, 1000 or 1500 consecutive nucleotides of a nucleic acid according to the invention.

**[0041]** As used herein, an "isolated nucleic acid fragment" is a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. An isolated nucleic acid fragment in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA.

**[0042]** A "gene" refers to an assembly of nucleotides that encode a polypeptide, and includes cDNA and genomic DNA nucleic acids. "Gene" also refers to a nucleic acid fragment that expresses a specific protein or polypeptide, including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. "Native gene" refers to a gene as found in nature with its own regulatory sequences. "Chimeric gene" refers to any gene that is not a native gene, comprising regulatory and/or coding sequences that are not found together in nature. Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature. A chimeric gene may comprise coding sequences derived from different sources and/or regulatory sequences derived from different sources. "Endogenous gene" refers to a native gene in its natural location in the genome of an organism A "foreign" gene or "heterologous" gene refers to a gene not normally found in the host organism, but that is introduced into the host organism by gene transfer. Foreign genes can comprise native genes inserted into a non-native organism, or chimeric genes. A "transgene" is a gene that has been introduced into the genome by a transformation procedure.

**[0043]** "Heterologous" DNA refers to DNA not naturally located in the cell, or in a chromosomal site of the cell. Preferably, the heterologous DNA includes a gene foreign to the cell.

**[0044]** The term "genome" includes chromosomal as well as mitochondrial, chloroplast and viral DNA or RNA.

**[0045]** A nucleic acid molecule is "hybridizable" to another nucleic acid molecule, such as a cDNA, genomic DNA, or RNA, when a single stranded form of the nucleic acid molecule can anneal to the other nucleic acid molecule under the appropriate conditions of temperature and solution ionic strength (see *Sambrook et al.*, 1989 *infra*). Hybridization and washing conditions are well known and exemplified in Sambrook, J., Fritsch, E. F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (1989), particularly Chapter 11 and Table 11.1 therein (entirely incorporated herein by reference). The conditions of temperature and ionic strength determine the "stringency" of the hybridization.

**[0046]** Stringency conditions can be adjusted to screen for moderately similar fragments, such as homologous sequences from distantly related organisms, to highly similar fragments, such as genes that duplicate functional enzymes from closely related organisms. For preliminary screening for homologous nucleic acids, low stringency hybridization conditions, corresponding to a  $T_m$  of 55°, can be used, e.g., 5x SSC, 0.1% SDS, 0.25% milk, and no formamide; or 30% formamide, 5x SSC, 0.5% SDS. Moderate stringency hybridization conditions correspond to a higher  $T_m$ , e.g., 40% formamide, with 5x or 6x SSC. High stringency hybridization conditions correspond to the highest  $T_m$ , e.g., 50% formamide, 5x or 6x SSC. Hybridization requires that the two nucleic acids contain complementary sequences, although depending on the stringency of the hybridization, mismatches between bases are possible.

**[0047]** The term "complementary" is used to describe the relationship between nucleotide bases that are capable of hybridizing to one another. For example, with respect to DNA, adenosine is complementary to thymine and cytosine is complementary to guanine. Accordingly, the instant invention also includes isolated nucleic acid fragments that are complementary to the complete sequences as disclosed or used herein as well as those substantially similar nucleic acid sequences.

**[0048]** In a specific embodiment, the term "standard hybridization conditions" refers to a  $T_m$  of 55°C, and utilizes conditions as set forth above. In a preferred embodiment, the  $T_m$  is 60°C; in a more preferred embodiment, the  $T_m$  is 65°C.

**[0049]** Post-hybridization washes also determine stringency conditions. One set of preferred conditions uses a series of washes starting with 6X SSC, 0.5% SDS at room temperature for 15 minutes (min), then repeated with 2X SSC, 0.5% SDS at 45°C for 30 minutes, and then repeated twice with 0.2X SSC, 0.5% SDS at 50°C for 30 minutes. A more preferred set of stringent conditions uses higher temperatures in which the washes are identical to those above except for the temperature of the final two 30 min washes in 0.2X SSC, 0.5% SDS was increased to 60°C. Another preferred set of highly stringent conditions uses two final washes in 0.1X SSC, 0.1% SDS at 65°C. Hybridization requires that the two nucleic acids comprise complementary sequences, although depending on the stringency of the hybridization, mismatches between bases are possible.

**[0050]** The appropriate stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementation, variables well known in the art. The greater the degree of similarity or homology between two nucleotide sequences, the greater the value of  $T_m$  for hybrids of nucleic acids having those sequences. The relative stability (corresponding to higher  $T_m$ ) of nucleic acid hybridizations decreases in the following order: RNA:RNA, DNA:RNA, DNA:DNA. For hybrids of greater than 100 nucleotides in length, equations for calculating  $T_m$  have been derived (see Sambrook *et al.*, *supra*, 9.50-0.51). For hybridization with shorter nucleic acids, i.e., oligonucleotides, the position of mismatches becomes more important, and the length of the oligonucleotide determines its specificity (see Sambrook *et al.*, *supra*, 11.7-11.8).

**[0051]** In one embodiment the length for a hybridizable nucleic acid is at least about 10 nucleotides. Preferable a minimum length for a hybridizable nucleic acid is at least about 15 nucleotides; more preferably at least about 20 nucleotides; and most preferably the length is at least 30 nucleotides. Furthermore, the skilled artisan will recognize that the temperature and wash solution salt concentration may be adjusted as necessary according to factors such as length of the probe.

**[0052]** The term "probe" refers to a single-stranded nucleic acid molecule that can base pair with a complementary single stranded target nucleic acid to form a double-stranded molecule.

**[0053]** As used herein, the term "oligonucleotide" refers to a nucleic acid, generally of at least 18 nucleotides, that is hybridizable to a genomic DNA molecule, a cDNA molecule, a plasmid DNA or an mRNA molecule. Oligonucleotides can be labeled, e.g., with  $^{32}P$ -nucleotides or nucleotides to which a label, such as biotin, has been covalently conjugated. A labeled oligonucleotide can be used as a probe to detect the presence of a nucleic acid. Oligonucleotides (one or both of which may be labeled) can be used as PCR primers, either for cloning full length or a fragment of a nucleic acid, or to detect the presence of a nucleic acid. An oligonucleotide can also be used to form a triple helix with a DNA molecule. Generally, oligonucleotides are prepared synthetically, preferably on a nucleic acid synthesizer. Accordingly, oligonucleotides can be prepared with non-naturally occurring phosphoester analog bonds, such as thioester bonds, etc.

**[0054]** A "primer" is an oligonucleotide that hybridizes to a target nucleic acid sequence to create a double stranded nucleic acid region that can serve as an initiation point for DNA synthesis under suitable conditions. Such primers may be used in a polymerase chain reaction.

**[0055]** "Polymerase chain reaction" is abbreviated PCR and means an *in vitro* method for enzymatically amplifying specific nucleic acid sequences. PCR involves a repetitive series of temperature cycles with each cycle comprising three



stages: denaturation of the template nucleic acid to separate the strands of the target molecule, annealing a single stranded PCR oligonucleotide primer to the template nucleic acid, and extension of the annealed primer(s) by DNA polymerase. PCR provides a means to detect the presence of the target molecule and, under quantitative or semi-quantitative conditions, to determine the relative amount of that target molecule within the starting pool of nucleic acids.

**[0056]** "Reverse transcription-polymerase chain reaction" is abbreviated RT-PCR and means an *in vitro* method for enzymatically producing a target cDNA molecule or molecules from an RNA molecule or molecules, followed by enzymatic amplification of a specific nucleic acid sequence or sequences within the target cDNA molecule or molecules as described above. RT-PCR also provides a means to detect the presence of the target molecule and, under quantitative or semi-quantitative conditions, to determine the relative amount of that target molecule within the starting pool of nucleic acids.

**[0057]** A DNA "coding sequence" is a double-stranded DNA sequence that is transcribed and translated into a polypeptide in a cell *in vitro* or *in vivo* when placed under the control of appropriate regulatory sequences. "Suitable regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include promoters, translation leader sequences, introns, polyadenylation recognition sequences, RNA processing site, effector binding site and stem-loop structure. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxyl) terminus. A coding sequence can include, but is not limited to, prokaryotic sequences, cDNA from mRNA, genomic DNA sequences, and even synthetic DNA sequences. If the coding sequence is intended for expression in a eukaryotic cell, a polyadenylation signal and transcription termination sequence will usually be located 3' to the coding sequence.

**[0058]** "Open reading frame" is abbreviated ORF and means a length of nucleic acid sequence, either DNA, cDNA or RNA, that comprises a translation start signal or initiation codon, such as an ATG or AUG, and a termination codon and can be potentially translated into a polypeptide sequence.

**[0059]** The term "head-to-head" is used herein to describe the orientation of two polynucleotide sequences in relation to each other. Two polynucleotides are positioned in a head-to-head orientation when the 5' end of the coding strand of one polynucleotide is adjacent to the 5' end of the coding strand of the other polynucleotide, whereby the direction of transcription of each polynucleotide proceeds away from the 5' end of the other polynucleotide. The term "head-to-head" may be abbreviated (5')-to-(5') and may also be indicated by the symbols ( $\leftarrow \rightarrow$ ) or (3' $\leftarrow$ 5' $\rightarrow$ 3').

**[0060]** The term "tail-to-tail" is used herein to describe the orientation of two polynucleotide sequences in relation to each other. Two polynucleotides are positioned in a tail-to-tail orientation when the 3' end of the coding strand of one polynucleotide is adjacent to the 3' end of the coding strand of the other polynucleotide, whereby the direction of transcription of each polynucleotide proceeds toward the other polynucleotide. The term "tail-to-tail" may be abbreviated (3')-to-(3') and may also be indicated by the symbols ( $\rightarrow \leftarrow$ ) or (5' $\rightarrow$ 3'3' $\leftarrow$ 5').

**[0061]** The term "head-to-tail" is used herein to describe the orientation of two polynucleotide sequences in relation to each other. Two polynucleotides are positioned in a head-to-tail orientation when the 5' end of the coding strand of one polynucleotide is adjacent to the 3' end of the coding strand of the other polynucleotide, whereby the direction of transcription of each polynucleotide proceeds in the same direction as that of the other polynucleotide. The term "head-to-tail" may be abbreviated (5')-to-(3') and may also be indicated by the symbols ( $\rightarrow \rightarrow$ ) or (5' $\rightarrow$ 3'5' $\rightarrow$ 3').

**[0062]** The term "downstream" refers to a nucleotide sequence that is located 3' to reference nucleotide sequence. In particular, downstream nucleotide sequences generally relate to sequences that follow the starting point of transcription. For example, the translation initiation codon of a gene is located downstream of the start site of transcription.

**[0063]** The term "upstream" refers to a nucleotide sequence that is located 5' to reference nucleotide sequence. In particular, upstream nucleotide sequences generally relate to sequences that are located on the 5' side of a coding sequence or starting point of transcription. For example, most promoters are located upstream of the start site of transcription.

**[0064]** The terms "restriction endonuclease" and "restriction enzyme" refer to an enzyme that binds and cuts within a specific nucleotide sequence within double stranded DNA.

**[0065]** "Homologous recombination" refers to the insertion of a foreign DNA sequence into another DNA molecule, e.g., insertion of a vector in a chromosome. Preferably, the vector targets a specific chromosomal site for homologous recombination. For specific homologous recombination, the vector will contain sufficiently long regions of homology to sequences of the chromosome to allow complementary binding and incorporation of the vector into the chromosome. Longer regions of homology, and greater degrees of sequence similarity, may increase the efficiency of homologous recombination.

**[0066]** Several methods known in the art may be used to propagate a polynucleotide according to the invention. Once a suitable host system and growth conditions are established, recombinant expression vectors can be propagated and prepared in quantity. As described herein, the expression vectors which can be used include, but are not limited to, the following vectors or their derivatives: human or animal viruses such as vaccinia virus or adenovirus; insect viruses such as baculovirus; yeast vectors; bacteriophage vectors (e.g., lambda), and plasmid and cosmid DNA vectors, to name but

a few.

[0067] A "vector" is any means for the cloning of and/or transfer of a nucleic acid into a host cell. A vector may be a replicon to which another DNA segment may be attached so as to bring about the replication of the attached segment. A "replicon" is any genetic element (e.g., plasmid, phage, cosmid, chromosome, virus) that functions as an autonomous unit of DNA replication *in vivo*, i.e., capable of replication under its own control. The term "vector" includes both viral and nonviral means for introducing the nucleic acid into a cell *in vitro*, *ex vivo* or *in vivo*. A large number of vectors known in the art may be used to manipulate nucleic acids, incorporate response elements and promoters into genes, etc. Possible vectors include, for example, plasmids or modified viruses including, for example bacteriophages such as lambda derivatives, or plasmids such as PBR322 or pUC plasmid derivatives, or the Bluescript vector. For example, the insertion of the DNA fragments corresponding to response elements and promoters into a suitable vector can be accomplished by ligating the appropriate DNA - fragments into a chosen vector that has complementary cohesive termini. Alternatively, the ends of the DNA molecules may be enzymatically modified or any site may be produced by ligating nucleotide sequences (linkers) into the DNA termini. Such vectors may be engineered to contain selectable marker genes that provide for the selection of cells that have incorporated the marker into the cellular genome. Such markers allow identification and/or selection of host cells that incorporate and express the proteins encoded by the marker.

[0068] Viral vectors, and particularly retroviral vectors, have been used in a wide variety of gene delivery applications in cells, as well as living animal subjects. Viral vectors that can be used include but are not limited to retrovirus, adeno-associated virus, pox, baculovirus, vaccinia, herpes simplex, Epstein-Barr, adenovirus, geminivirus, and caulimovirus vectors. Non-viral vectors include plasmids, liposomes, electrically charged lipids (cytofectins), DNA-protein complexes, and biopolymers. In addition to a nucleic acid, a vector may also comprise one or more regulatory regions, and/or selectable markers useful in selecting, measuring, and monitoring nucleic acid transfer results (transfer to which tissues, duration of expression, etc.).

[0069] The term "plasmid" refers to an extra chromosomal element often carrying a gene that is not part of the central metabolism of the cell, and usually in the form of circular double-stranded DNA molecules. Such elements may be autonomously replicating sequences, genome integrating sequences, phage or nucleotide sequences, linear, circular, or supercoiled, of a single- or double-stranded DNA or RNA, derived from any source, in which a number of nucleotide sequences have been joined or recombined into a unique construction which is capable of introducing a promoter fragment and DNA sequence for a selected gene product along with appropriate 3' untranslated sequence into a cell.

[0070] A "cloning vector" is a "replicon", which is a unit length of a nucleic acid, preferably DNA, that replicates sequentially and which comprises an origin of replication, such as a plasmid, phage or cosmid, to which another nucleic acid segment may be attached so as to bring about the replication of the attached segment. Cloning vectors may be capable of replication in one cell type and expression in another ("shuttle vector").

[0071] Vectors may be introduced into the desired host cells by methods known in the art, e.g., transfection, electroporation, microinjection, transduction, cell fusion, DEAE dextran, calcium phosphate precipitation, lipofection (lysosome fusion), use of a gene gun, or a DNA vector transporter (see, e.g., Wu et al., 1992, J. Biol. Chem. 267:963-967; Wu and Wu, 1988, J. Biol. Chem. 263:14621-14624; and Hartmut et al., Canadian Patent Application No. 2,012,311, filed March 15, 1990).

[0072] A polynucleotide according to the invention can also be introduced *in vivo* by lipofection. For the past decade, there has been increasing use of liposomes for encapsulation and transfection of nucleic acids *in vitro*. Synthetic cationic lipids designed to limit the difficulties and dangers encountered with liposome mediated transfection can be used to prepare liposomes for *in vivo* transfection of a gene encoding a marker (Felper et al., 1987, PNAS 84:7413; Mackey, et al., 1988, Proc. Natl. Acad. Sci. U.S.A. 85:8027-8031; and Ulmer et al., 1993, Science 259:1745-1748). The use of cationic lipids may promote encapsulation of negatively charged nucleic acids, and also promote fusion with negatively charged cell membranes (Felgner and Ringold, 1989, Science 337:387-388). Particularly useful lipid compounds and compositions for transfer of nucleic acids are described in International Patent Publications WO95/18863 and WO96/17823, and in U.S. Patent No. 5,459,127. The use of lipofection to introduce exogenous genes into the specific organs *in vivo* has certain practical advantages. Molecular targeting of liposomes to specific cells represents one area of benefit. It is clear that directing transfection to particular cell types would be particularly preferred in a tissue with cellular heterogeneity, such as pancreas, liver, kidney, and the brain. Lipids may be chemically coupled to other molecules for the purpose of targeting (Mackey, et al., 1988, *supra*). Targeted peptides, e.g., hormones or neurotransmitters, and proteins such as antibodies, or non-peptide molecules could be coupled to liposomes chemically.

[0073] Other molecules are also useful for facilitating transfection of a nucleic acid *in vivo*, such as a cationic oligopeptide (e.g., WO95/21931), peptides derived from DNA binding proteins (e.g., WO96/25508), or a cationic polymer (e.g., WO95/21931).

[0074] It is also possible to introduce a vector *in vivo* as a naked DNA plasmid (see U.S. Patents 5,693,622, 5,589,466 and 5,580,859). Receptor-mediated DNA delivery approaches can also be used (Curiel et al., 1992, Hum. Gene Ther. 3:147-154; and Wu and Wu, 1987, J. Biol. Chem 262:4429-4432).

[0075] The term "transfection" means the uptake of exogenous or heterologous RNA or DNA by a cell. A cell has been

"transfected" by exogenous or heterologous RNA or DNA when such RNA or DNA has been introduced inside the cell. A cell has been "transformed" by exogenous or heterologous RNA or DNA when the transfected RNA or DNA effects a phenotypic change. The transforming RNA or DNA can be integrated (covalently linked) into chromosomal DNA making up the genome of the cell.

**[0076]** "Transformation" refers to the transfer of a nucleic acid fragment into the genome of a host organism, resulting in genetically stable inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as "transgenic" or "recombinant" or "transformed" organisms.

**[0077]** The term "genetic region" will refer to a region of a nucleic acid molecule or a nucleotide sequence that comprises a gene encoding a polypeptide.

**[0078]** In addition, the recombinant vector comprising a polynucleotide according to the invention may include one or more origins for replication in the cellular hosts in which their amplification or their expression is sought, markers or selectable markers.

**[0079]** The term "selectable marker" means an identifying factor, usually an antibiotic or chemical resistance gene, that is able to be selected for based upon the marker gene's effect, i.e., resistance to an antibiotic, resistance to a herbicide, colorimetric markers, enzymes, fluorescent markers, and the like, wherein the effect is used to track the inheritance of a nucleic acid of interest and/or to identify a cell or organism that has inherited the nucleic acid of interest. Examples of selectable marker genes known and used in the art include: genes providing resistance to ampicillin, streptomycin, gentamycin, kanamycin, hygromycin, bialaphos herbicide, sulfonamide, and the like; and genes that are used as phenotypic markers, i.e., anthocyanin regulatory genes, isopentenyl transferase gene, and the like.

**[0080]** The term "reporter gene" means a nucleic acid encoding an identifying factor that is able to be identified based upon the reporter gene's effect, wherein the effect is used to track the inheritance of a nucleic acid of interest, to identify a cell or organism that has inherited the nucleic acid of interest, and/or to measure gene expression induction or transcription. Examples of reporter genes known and used in the art include: luciferase (Luc), green fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT),  $\beta$ -galactosidase (LacZ),  $\beta$ -glucuronidase (Gus), and the like. Selectable marker genes may also be considered reporter genes.

**[0081]** "Promoter" refers to a DNA sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding sequence is located 3' to a promoter sequence. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic DNA segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental or physiological conditions. Promoters that cause a gene to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". Promoters that cause a gene to be expressed in a specific cell type are commonly referred to as "cell-specific promoters" or "tissue-specific promoters". Promoters that cause a gene to be expressed at a specific stage of development or cell differentiation are commonly referred to as "developmentally-specific promoters" or "cell differentiation-specific promoters". Promoters that are induced and cause a gene to be expressed following exposure or treatment of the cell with an agent, biological molecule, chemical, ligand, light, or the like that induces the promoter are commonly referred to as "inducible promoters" or "regulatable promoters". It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of different lengths may have identical promoter activity.

**[0082]** A "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site (conveniently defined for example, by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase.

**[0083]** A coding sequence is "under the control" of transcriptional and translational control sequences in a cell when RNA polymerase transcribes the coding sequence into mRNA, which is then trans-RNA spliced (if the coding sequence contains introns) and translated into the protein encoded by the coding sequence.

**[0084]** "Transcriptional and translational control sequences" are DNA regulatory sequences, such as promoters, enhancers, terminators, and the like, that provide for the expression of a coding sequence in a host cell. In eukaryotic cells, polyadenylation signals are control sequences.

**[0085]** The term "response element" means one or more cis-acting DNA elements which confer responsiveness on a promoter mediated through interaction with the DNA-binding domains of the first chimeric gene. This DNA element may be either palindromic (perfect or imperfect) in its sequence or composed of sequence motifs or half sites separated by a variable number of nucleotides. The half sites can be similar or identical and arranged as either direct or inverted repeats or as a single half site or multimers of adjacent half sites in tandem. The response element may comprise a minimal promoter isolated from different organisms depending upon the nature of the cell or organism into which the

response element will be incorporated. The DNA binding domain of the first hybrid protein binds, in the presence or absence of a ligand, to the DNA sequence of a response element to initiate or suppress transcription of downstream gene(s) under the regulation of this response element. Examples of DNA sequences for response elements of the natural ecdysone receptor include: RRGG/TTCANTGAC/ACY (see Cherbas L., et al., (1991), *Genes Dev.* 5, 120-131); AG-GTCAN<sub>10</sub>AGGTCA, where N<sub>10</sub> can be one or more spacer nucleotides (see D'Avino PP., et al., (1995), *Mol. Cell. Endocrinol.* 113, 1-9); and GGGTTGAATGAATTT (see Antoniewski C., et al., (1994), *Mol. Cell Biol.* 14, 4465-4474).

**[0086]** The term "operably linked" refers to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable of affecting the expression of that coding sequence (i.e., that the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

**[0087]** The term "expression", as used herein, refers to the transcription and stable accumulation of sense (mRNA) or antisense RNA derived from a nucleic acid or polynucleotide. Expression may also refer to translation of mRNA into a protein or polypeptide.

**[0088]** The terms "cassette", "expression cassette" and "gene expression cassette" refer to a segment of DNA that can be inserted into a nucleic acid or polynucleotide at specific restriction sites or by homologous recombination. The segment of DNA comprises a polynucleotide that encodes a polypeptide of interest, and the cassette and restriction sites are designed to ensure insertion of the cassette in the proper reading frame for transcription and translation. "Transformation cassette" refers to a specific vector comprising a polynucleotide that encodes a polypeptide of interest and having elements in addition to the polynucleotide that facilitate transformation of a particular host cell. Cassettes, expression cassettes, gene expression cassettes and transformation cassettes of the invention may also comprise elements that allow for enhanced expression of a polynucleotide encoding a polypeptide of interest in a host cell. These elements may include, but are not limited to: a promoter, a minimal promoter, an enhancer, a response element, a terminator sequence, a polyadenylation sequence, and the like.

**[0089]** For purposes of this invention, the term "gene switch" refers to the combination of a response element associated with a promoter, and an ECR based system which, in the presence of one or more ligands, modulates the expression of a gene into which the response element and promoter are incorporated.

**[0090]** The terms "modulate" and "modulates" mean to induce, reduce or inhibit nucleic acid or gene expression, resulting in the respective induction, reduction or inhibition of protein or polypeptide production.

**[0091]** The plasmids or vectors according to the invention may further comprise at least one promoter suitable for driving expression of a gene in a host cell. The term "expression vector" means a vector, plasmid or vehicle designed to enable the expression of an inserted nucleic acid sequence following transformation into the host. The cloned gene, i.e., the inserted nucleic acid sequence, is usually placed under the control of control elements such as a promoter, a minimal promoter, an enhancer, or the like. Initiation control regions or promoters, which are useful to drive expression of a nucleic acid in the desired host cell are numerous and familiar to those skilled in the art. Virtually any promoter capable of driving these genes is suitable for the present invention including but not limited to: viral promoters, plant promoters, bacterial promoters, animal promoters, mammalian promoters, synthetic promoters, constitutive promoters, tissue specific promoter, developmental specific promoters, inducible promoters, light regulated promoters: *CYC1*, *HIS3*, *GAL1*, *GAL4*, *GAL10*, *ADHI*, *PGK*, *PHO5*, *GAPDH*, *ADC1*, *TRP1*, *URA3*, *LEU2*, *ENO*, *TPI*, alkaline phosphatase promoters (useful for expression in *Saccharomyces*): *AOX1* promoter (useful for expression in *Pichia*); b-lactamase, *lac*, *ara*, *tet*, *trp*, *IP<sub>4</sub>*, *IP<sub>3</sub>*, *T7*, *tac*, and *trc* promoters (useful for expression in *Escherichia coli*); and light regulated-, seed specific-, pollen specific-, ovary specific-, pathogenesis or disease related-, cauliflower mosaic virus 35S, CMV 35S minimal, cassava vein mosaic virus (CsVMV), chlorophyll a/b binding protein, ribulose 1, 5-bisphosphate carboxylase, shoot-specific, root specific, chitinase, stress inducible, rice tungro bacilliform virus, plant super-promoter, potato leucine aminopeptidase, nitrate reductase, mannopine synthase, nopaline synthase, ubiquitin, zein protein, and anthocyanin promoters (useful for expression in plant cells); animal and mammalian promoters known in the art include, but are not limited to, the SV40 early (SV40e) promoter region, the promoter contained in the 3' long terminal repeat (LTR) of Rous sarcoma virus (RSV), the promoters of the E1A or major late promoter (MLP) genes of adenoviruses, the cytomegalovirus early promoter, the herpes simplex virus (HSV) thymidine kinase (TK) promoter, an elongation factor 1 alpha (EF1) promoter, a phosphoglycerate kinase (PGK) promoter, a ubiquitin (Ubc) promoter, an albumin promoter, the regulatory sequences of the mouse metallothionein-L promoter, and transcriptional control regions, the ubiquitous promoters (HPRT, vimentin,  $\alpha$ -actin, tubulin and the like), the promoters of the intermediate filaments (desmin, neuro-filaments, keratin, GFAP, and the like), the promoters of therapeutic genes (of the MDR, CFTR or factor VIII type, and the like), and promoters that exhibit tissue specificity and have been utilized in transgenic animals, such as the elastase I gene control region which is active in pancreatic acinar cells; insulin gene control region active in pancreatic beta cells, immunoglobulin gene control region active in lymphoid cells, mouse mammary tumor virus control region active in testicular, breast, lymphoid and mast cells; albumin gene, Apo AI and Apo AII control regions active in liver, alpha-fetoprotein gene control region active in liver, alpha 1-antitrypsin gene control region active in the liver, beta-globin gene

control region active in myeloid cells, myelin basic protein gene control region ) active in oligodendrocyte cells in the brain, myosin light chain-2 gene control region active in skeletal muscle, and gonadotropin releasing hormone gene control region active in the hypothalamus, pyruvate kinase promoter, villin promoter, promoter of the fatty acid binding intestinal protein, promoter of the smooth muscle cell  $\alpha$ -actin, and the like. In a preferred embodiment of the invention, the promoter is selected from the group consisting of a cauliflower mosaic virus 35S promoter, a cassava vein mosaic virus promoter, and a cauliflower mosaic virus 35S minimal promoter, an elongation factor 1 alpha (EF1) promoter, a phosphoglycerate kinase (PGK) promoter, a ubiquitin (Ubc) promoter, and an albumin promoter. In addition, these expression sequences may be modified by addition of enhancer or regulatory sequences and the like.

**[0092]** Enhancers that may be used in embodiments of the invention include but are not limited to: tobacco mosaic virus enhancer, cauliflower mosaic virus 35S enhancer, tobacco etch virus enhancer, ribulose 1, 5-bisphosphate carboxylase enhancer, rice tungro bacilliform virus enhancer, and other plant and viral gene enhancers, and the like.

**[0093]** Termination control regions, *i.e.*, terminator or polyadenylation sequences, may also be derived from various genes native to the preferred hosts. Optionally, a termination site may be unnecessary, however, it is most preferred if included. In a preferred embodiment of the invention, the termination control region may be comprised or be derived from a synthetic sequence, synthetic polyadenylation signal, an SV40 late polyadenylation signal, an SV40 polyadenylation signal, a bovine growth hormone (BGH) polyadenylation signal, nopaline synthase (nos), cauliflower mosaic virus (CaMV), octopine synthase (ocs), Agrobacterium, viral, and plant terminator sequences, or the like.

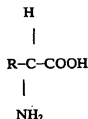
**[0094]** The terms "3' non-coding sequences" or "3' untranslated region (UTR)" refer to DNA sequences located downstream (3') of a coding sequence and may comprise polyadenylation [poly(A)] recognition sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor.

**[0095]** "Regulatory region" means a nucleic acid sequence which regulates the expression of a second nucleic acid sequence. A regulatory region may include sequences which are naturally responsible for expressing a particular nucleic acid (a homologous region) or may include sequences of a different origin that are responsible for expressing different proteins or even synthetic proteins (a heterologous region). In particular, the sequences can be sequences of prokaryotic, eukaryotic, or viral genes or derived sequences that stimulate or repress transcription of a gene in a specific or non-specific manner and in an inducible or non-inducible manner. Regulatory regions include origins of replication, RNA splice sites, promoters, enhancers, transcriptional termination sequences, and signal sequences which direct the polypeptide into the secretory pathways of the target cell.

**[0096]** A regulatory region from a "heterologous source" is a regulatory region that is not naturally associated with the expressed nucleic acid. Included among the heterologous regulatory regions are regulatory regions from a different species, regulatory regions from a different gene, hybrid regulatory sequences, and regulatory sequences which do not occur in nature, but which are designed by one having ordinary skill in the art.

**[0097]** "RNA transcript" refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be a RNA sequence derived from post-transcriptional processing of the primary transcript and is referred to as the mature RNA. "Messenger RNA (mRNA)" refers to the RNA that is without introns and that can be translated into protein by the cell. "cDNA" refers to a double-stranded DNA that is complementary to and derived from mRNA. "Sense" RNA refers to RNA transcript that includes the mRNA and so can be translated into protein by the cell. "Antisense RNA" refers to a RNA transcript that is complementary to all or part of a target primary transcript or mRNA and that blocks the expression of a target gene. The complementarity of an antisense RNA may be with any part of the specific gene transcript, *i.e.*, at the 5' non-coding sequence, 3' non-coding sequence, or the coding sequence. "Functional RNA" refers to antisense RNA, ribozyme RNA, or other RNA that is not translated yet has an effect on cellular processes.

**[0098]** A "polypeptide" is a polymeric compound comprised of covalently linked amino acid residues. Amino acids have the following general structure:



Amino acids are classified into seven groups on the basis of the side chain R: (1) aliphatic side chains, (2) side chains containing a hydroxylic (OH) group, (3) side chains containing sulfur atoms, (4) side chains containing an acidic or amide

group, (5) side chains containing a basic group, (6) side chains containing an aromatic ring, and (7) proline, an imino acid in which the side chain is fused to the amino group. A polypeptide of the invention preferably comprises at least about 14 amino acids.

[0099] A "protein" is a polypeptide that performs a structural or functional role in a living cell.

[0100] An "isolated polypeptide" or "isolated protein" is a polypeptide or protein that is substantially free of those compounds that are normally associated therewith in its natural state (e.g., other proteins or polypeptides, nucleic acids, carbohydrates, lipids). "Isolated" is not meant to exclude artificial or synthetic mixtures with other compounds, or the presence of impurities which do not interfere with biological activity, and which may be present, for example, due to incomplete purification, addition of stabilizers, or compounding into a pharmaceutically acceptable preparation.

[0101] "Fragment" of a polypeptide according to the invention will be understood to mean a polypeptide whose amino acid sequence is shorter than that of the reference polypeptide and which comprises, over the entire portion with these reference polypeptides, an identical amino acid sequence. Such fragments may, where appropriate, be included in a larger polypeptide of which they are a part. Such fragments of a polypeptide according to the invention may have a length of 10, 15, 20, 30 to 40, 50, 100, 200 or 300 amino acids.

[0102] A "variant" of a polypeptide or protein is any analogue, fragment, derivative, or mutant which is derived from a polypeptide or protein and which retains at least one biological property of the polypeptide or protein. Different variants of the polypeptide or protein may exist in nature. These variants may be allelic variations characterized by differences in the nucleotide sequences of the structural gene coding for the protein, or may involve differential splicing or post-translational modification. The skilled artisan can produce variants having single or multiple amino acid substitutions, deletions, additions, or replacements. These variants may include, *inter alia*: (a) variants in which one or more amino acid residues are substituted with conservative or non-conservative amino acids, (b) variants in which one or more amino acids are added to the polypeptide or protein, (c) variants in which one or more of the amino acids includes a substituent group, and (d) variants in which the polypeptide or protein is fused with another polypeptide such as serum albumin. The techniques for obtaining these variants, including genetic (suppressions, deletions, mutations, etc.), chemical, and enzymatic techniques, are known to persons having ordinary skill in the art. A variant polypeptide preferably comprises at least about 14 amino acids.

[0103] A "heterologous protein" refers to a protein not naturally produced in the cell.

[0104] A "mature protein" refers to a post-translationally processed polypeptide; i.e., one from which any pre- or propeptides present in the primary translation product have been removed. "Precursor" protein refers to the primary product of translation of mRNA; i.e., with pre- and propeptides still present. Pre- and propeptides may be but are not limited to intracellular localization signals.

[0105] The term "signal peptide" refers to an amino terminal polypeptide preceding the secreted mature protein. The signal peptide is cleaved from and is therefore not present in the mature protein. Signal peptides have the function of directing and translocating secreted proteins across cell membranes. Signal peptide is also referred to as signal protein.

[0106] A "signal sequence" is included at the beginning of the coding sequence of a protein to be expressed on the surface of a cell. This sequence encodes a signal peptide, N-terminal to the mature polypeptide, that directs the host cell to translocate the polypeptide. The term "translocation signal sequence" is used herein to refer to this sort of signal sequence. Translocation signal sequences can be found associated with a variety of proteins native to eukaryotes and prokaryotes, and are often functional in both types of organisms.

[0107] The term "homology" refers to the percent of identity between two polynucleotide or two polypeptide moieties. The correspondence between the sequence from one moiety to another can be determined by techniques known to the art. For example, homology can be determined by a direct comparison of the sequence information between two polypeptide molecules by aligning the sequence information and using readily available computer programs. Alternatively, homology can be determined by hybridization of polynucleotides under conditions that form stable duplexes between homologous regions, followed by digestion with single-stranded-specific nuclease(s) and size determination of the digested fragments.

[0108] As used herein, the term "homologous" in all its grammatical forms and spelling variations refers to the relationship between proteins that possess a "common evolutionary origin," including proteins from superfamilies (e.g., the immunoglobulin superfamily) and homologous proteins from different species (e.g., myosin light chain, etc.) (Reeck et al., 1987, Cell 50:667.). Such proteins (and their encoding genes) have sequence homology, as reflected by their high degree of sequence similarity.

[0109] Accordingly, the term "sequence similarity" in all its grammatical forms refers to the degree of identity or correspondence between nucleic acid or amino acid sequences of proteins that may or may not share a common evolutionary origin (see Reeck et al., 1987, Cell 50:667). As used herein, the term "homologous" in all its grammatical forms and spelling variations refers to the relationship between proteins that possess a "common evolutionary origin," including proteins from superfamilies and homologous proteins from different species (Reeck et al., *supra*). Such proteins (and their encoding genes) have sequence homology, as reflected by their high degree of sequence similarity. However, in common usage and in the instant application, the term "homologous," when modified with an adverb such as "highly,"

may refer to sequence similarity and not a common evolutionary origin.

**[0110]** In a specific embodiment, two DNA sequences are "substantially homologous" or "substantially similar" when at least about 50% (preferably at least about 75%, and most preferably at least about 90 or 95%) of the nucleotides match over the defined length of the DNA sequences. Sequences that are substantially homologous can be identified by comparing the sequences using standard software available in sequence data banks, or in a Southern hybridization experiment under, for example, stringent conditions as defined for that particular system. Defining appropriate hybridization conditions is within the skill of the art. See, e.g., Sambrook *et al.*, 1989, *supra*.

**[0111]** As used herein, "substantially similar" refers to nucleic acid fragments wherein changes in one or more nucleotide bases results in substitution of one or more amino acids, but do not affect the functional properties of the protein encoded by the DNA sequence. "Substantially similar" also refers to nucleic acid fragments wherein changes in one or more nucleotide bases do not affect the ability of the nucleic acid fragment to mediate alteration of gene expression by antisense or co-suppression technology. "Substantially similar" also refers to modifications of the nucleic acid fragments of the instant invention such as deletion or insertion of one or more nucleotide bases that do not substantially affect the functional properties of the resulting transcript. It is therefore understood that the invention encompasses more than the specific exemplary sequences. Each of the proposed modifications is well within the routine skill in the art, as is determination of retention of biological activity of the encoded products.

**[0112]** Moreover, the skilled artisan recognizes that substantially similar sequences encompassed by this invention are also defined by their ability to hybridize, under stringent conditions (0.1X SSC, 0.1% SDS, 65 °C and washed with 2X SSC, 0.1% SDS followed by 0.1X SSC, 0.1% SDS), with the sequences exemplified herein. Substantially similar nucleic acid fragments of the instant invention are those nucleic acid fragments whose DNA sequences are at least 70% identical to the DNA sequence of the nucleic acid fragments reported herein. Preferred substantially nucleic acid fragments of the instant invention are those nucleic acid fragments whose DNA sequences are at least 80% identical to the DNA sequence of the nucleic acid fragments reported herein. More preferred nucleic acid fragments are at least 90% identical to the DNA sequence of the nucleic acid fragments reported herein. Even more preferred are nucleic acid fragments that are at least 95% identical to the DNA sequence of the nucleic acid fragments reported herein.

**[0113]** Two amino acid sequences are "substantially homologous" or "substantially similar" when greater than about 40% of the amino acids are identical, or greater than 60% are similar (functionally identical). Preferably, the similar or homologous sequences are identified by alignment using, for example, the GCG (Genetics Computer Group, Program Manual for the GCG Package, Version 7, Madison, Wisconsin) pileup program

**[0114]** The term "corresponding to" is used herein to refer to similar or homologous sequences, whether the exact position is identical or different from the molecule to which the similarity or homology is measured. A nucleic acid or amino acid sequence alignment may include spaces. Thus, the term "corresponding to" refers to the sequence similarity, and not the numbering of the amino acid residues or nucleotide bases.

**[0115]** A "substantial portion" of an amino acid or nucleotide sequence comprises enough of the amino acid sequence of a polypeptide or the nucleotide sequence of a gene to putatively identify that polypeptide or gene, either by manual evaluation of the sequence by one skilled in the art, or by computer-automated sequence comparison and identification using algorithms such as BLAST (Basic Local Alignment Search Tool; Altschul, S. F., *et al.*, (1993) *J. Mol. Biol.* 215: 403-410; see also [www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/)). In general, a sequence of ten or more contiguous amino acids or thirty or more nucleotides is necessary in order to putatively identify a polypeptide or nucleic acid sequence as homologous to a known protein or gene. Moreover, with respect to nucleotide sequences, gene specific oligonucleotide probes comprising 20-30 contiguous nucleotides may be used in sequence-dependent methods of gene identification (e.g., Southern hybridization) and isolation (e.g., *in situ* hybridization of bacterial colonies or bacteriophage plaques). In addition, short oligonucleotides of 12-15 bases may be used as amplification primers in PCR in order to obtain a particular nucleic acid fragment comprising the primers. Accordingly, a "substantial portion" of a nucleotide sequence comprises enough of the sequence to specifically identify and/or isolate a nucleic acid fragment comprising the sequence.

**[0116]** The term "percent identity", as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" and "similarity" can be readily calculated by known methods, including but not limited to those described in: *Computational Molecular Biology* (Lesk, A. M., ed.) Oxford University Press, New York (1988); *Bioinformatics: Informatics and Genome Projects* (Smith, D. W., ed.) Academic Press, New York (1993); *Computer Analysis of Sequence Data, Part I* (Griffin, A. M., and Griffin, H. G., eds.) Humana Press, New Jersey (1994); *Sequence Analysis in Molecular Biology* (von Heinje, G., ed.) Academic Press (1987); and *Sequence Analysis Primer* (Gribkov, M. and Devereux, J., eds.) Stockton Press, New York (1991). Preferred methods to determine identity are designed to give the best match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Sequence alignments and percent identity calculations may be performed using the Megalign program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences may be performed using the Clustal method of alignment (Higgins

and Sharp (1989) *CABIOS*, 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the Clustal method may be selected: KTDUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

[0117] The term "sequence analysis software" refers to any computer algorithm or software program that is useful for the analysis of nucleotide or amino acid sequences. "Sequence analysis software" may be commercially available or independently developed. Typical sequence analysis software will include but is not limited to the GCG suite of programs (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, WI), BLASTP, BLASTN, BLASTX (Altschul et al., *J. Mol. Biol.* 215:403-410 (1990)), and DNASTAR (DNASTAR, Inc. 1228 S. Park St. Madison, WI 53715 USA). Within the context of this application it will be understood that where sequence analysis software is used for analysis, that the results of the analysis will be based on the "default values" of the program referenced, unless otherwise specified. As used herein "default values" will mean any set of values or parameters which originally load with the software when first initialized.

[0118] "Synthetic genes" can be assembled from oligonucleotide building blocks that are chemically synthesized using procedures known to those skilled in the art. These building blocks are ligated and annealed to form gene segments that are then enzymatically assembled to construct the entire gene. "Chemically synthesized", as related to a sequence of DNA, means that the component nucleotides were assembled *in vitro*. Manual chemical synthesis of DNA may be accomplished using well established procedures, or automated chemical synthesis can be performed using one of a number of commercially available machines. Accordingly, the genes can be tailored for optimal gene expression based on optimization of nucleotide sequence to reflect the codon bias of the host cell. The skilled artisan appreciates the likelihood of successful gene expression if codon usage is biased towards those codons favored by the host. Determination of preferred codons can be based on a survey of genes derived from the host cell where sequence information is available.

#### **GENE EXPRESSION MODULATION SYSTEM OF THE INVENTION**

[0119] Applicants have now shown that separating the transactivation and DNA binding domains by placing them on two different proteins results in greatly reduced background activity in the absence of a ligand and significantly increased activity over background in the presence of a ligand. Applicants' improved gene expression system comprises two chimeric gene expression; the first encoding a DNA binding domain fused to a nuclear receptor polypeptide and the second encoding a transactivation domain fused to a nuclear receptor polypeptide. The interaction of the first protein with the second protein effectively tethers the DNA binding domain to the transactivation domain. Since the DNA binding and transactivation domains reside on two different molecules, the background activity in the absence of ligand is greatly reduced.

[0120] In general, the inducible gene expression modulation system of the invention comprises a) a first chimeric gene that is capable of being expressed in a host cell comprising a polynucleotide sequence that encodes a first hybrid polypeptide comprising i) a DNA binding domain that recognizes a response element associated with a gene whose expression is to be modulated; and ii) a ligand binding domain comprising the ligand binding domain from a nuclear receptor, and b) a second chimeric gene that is capable of being expressed in the host cell comprising a polynucleotide sequence that encodes a second hybrid polypeptide comprising: i) a transactivation domain; and ii) a ligand binding domain comprising the ligand binding domain from a nuclear receptor other than ultraspiracle (USP); wherein the transactivation domain are from other than EcR, RXR, or USP; and wherein the ligand binding domains from the first hybrid polypeptide and the second hybrid polypeptide are different and dimerize.

[0121] This two-hybrid system exploits the ability of a pair of interacting proteins to bring the transcription activation domain into a more favorable position relative to the DNA binding domain such that when the DNA binding domain binds to the DNA binding site on the gene, the transactivation domain more effectively activates the promoter (see, for example, U.S. Patent No. 5,283,173). This two-hybrid system is a significantly improved inducible gene expression modulation system compared to the two systems disclosed in International Patent Applications PCT/US97/05330 and PCT/US98/14215.

[0122] The ecodyson receptor-based gene expression modulation system of the invention may be either heterodimeric and homodimeric. A functional EcR complex generally refers to a heterodimeric protein complex consisting of two members of the steroid receptor family, an ecodyson receptor protein obtained from various insects, and an ultraspiracle (USP) protein or the vertebrate homolog of USP, retinoid X receptor protein (see Yao, et al. (1993) *Nature* 366, 476-479; Yao, et al., (1992) *Cell* 71, 63-72). However, the complex may also be a homodimer as detailed below. The functional ecodysteroid receptor complex may also include additional protein(s) such as immunophilins. Additional members of the steroid receptor family of proteins, known as transcriptional factors (such as DHR38 or *betaFTZ-1*), may also be ligand dependent or independent partners for EcR, USP, and/or RXR. Additionally, other cofactors may be required such as proteins generally known as coactivators (also termed adapters or mediators). These proteins do not bind sequence-specifically to DNA and are not involved in basal transcription. They may exert their effect on transcription activation



through various mechanisms, including stimulation of DNA-binding of activators, by affecting chromatin structure, or by mediating activator-initiation complex interactions. Examples of such coactivators include RIP140, TIF1, RAP46Bag-1, ARA70, SRC-1/NCoA-1, TIF2/GRIP/NCoA-2, ACTR/AIB 1/RAC3/pCIP as well as the promiscuous co activator C response element B binding protein, CBP/p300 (for review see Glass et al, *Curr. Opin. Cell Biol.* 9:222-232, 1997). Also, protein cofactors generally known as corepressors (also known as repressors, silencers, or silencing mediators) may be required to effectively inhibit transcriptional activation in the absence of ligand. These corepressors may interact with the unliganded ecdysone receptor to silence the activity at the response element. Current evidence suggests that binding of ligand changes the conformation of the receptor, which results in release of the corepressor and recruitment of the above described coactivators, thereby, abolishing their silencing activity. Examples of corepressors include N-CoR and SMRT (for review, see Horwitz et al. *Mol Endocrinol.* 10: 1167-1177, 1996). These cofactors may either be endogenous within the cell or organism, or may be added exogenously as transgenes to be expressed in either a regulated or unregulated fashion. Homodimer complexes of the ecdysone receptor protein, USP, or RXR may also be functional under some circumstances.

**[0123]** The ecdysone receptor complex typically includes proteins which are members of the nuclear receptor superfamily wherein all members are characterized by the presence of an amino-terminal transactivation domain, a DNA binding domain ("DBD"), and a ligand binding domain ("LBD") separated from the DBD by a hinge region. As used herein, the term "DNA binding domain" comprises a minimal peptide sequence of a DNA binding protein, up to the entire length of a DNA binding protein, so long as the DNA binding domain functions to associate with a particular response element. Members of the nuclear receptor superfamily are also characterized by the presence of four or five domains: A/B, C, D, E, and in some members F (see Evans, *Science* 240:889-895 (1988)). The "AB" domain corresponds to the transactivation domain, "C" corresponds to the DNA binding domain, "D" corresponds to the hinge region, and "E" corresponds to the ligand binding domain. Some members of the family may also have another transactivation domain on the carboxy-terminal side of the LBD corresponding to "F".

**[0124]** The DBD is characterized by the presence of two cysteine zinc fingers between which are two amino acid motifs, the P-box and the D-box, which confer specificity for ecdysone response elements. These domains may be either native, modified, or chimeras of different domains of heterologous receptor proteins. This EcR receptor, like a subset of the steroid receptor family, also possesses less well defined regions responsible for heterodimerization properties. Because the domains of EcR, USP, and RXR are modular in nature, the LBD, DBD, and transactivation domains may be interchangeable.

**[0125]** Gene switch systems are known that incorporate components from the ecdysone receptor complex. However, in these known systems, whenever EcR is used it is associated with native or modified DNA binding domains and transactivation domains on the same molecule. USP or RXR are typically used as silent partners. We have now shown that when DNA binding domains and transactivation domains are on the same molecule the background activity in the absence of ligand is high and that such activity is dramatically reduced when DNA binding domains and transactivation domains are on different molecules, that is, on each of two partners of a heterodimeric or homodimeric complex. This two-hybrid system also provides improved sensitivity to non-steroidal ligands for example, diacylhydrazines, when compared to steroidal ligands for example, ponasterone A ("PonA") or muristerone A ("MuRA"). That is, when compared to steroids, the non-steroidal ligands provide higher activity at a lower concentration. In addition, since transactivation based on EcR gene switches is often cell-line dependent, it is easier to tailor switching system to obtain maximum transactivation capability for each application. Furthermore, this two-hybrid system avoids some side effects due to overexpression of RXR that often occur when unmodified RXR is used as a switching partner. In this two-hybrid system, native DNA binding and transactivation domains of EcR or RXR are eliminated. As a result, these chimeric molecules have less chance of interacting with other steroid hormone receptors present in the cell resulting in reduced side effects.

**[0126]** Specifically, Applicants' invention relates to a gene expression modulation system comprising: a) a first gene expression cassette that is capable of being expressed in a host cell, wherein the first gene expression cassette comprises a polynucleotide that encodes a first polypeptide comprising i) a DNA-binding domain that recognizes a response element associated with a gene whose expression is to be modulated; and ii) a ligand binding domain comprising a ligand binding domain from a nuclear receptor; and b) a second gene expression cassette that is capable of being expressed in the host cell, wherein the second gene expression cassette comprises a polynucleotide sequence that encodes a second polypeptide comprising i) a transactivation domain; and ii) a ligand binding domain comprising a ligand binding domain from a nuclear receptor other than ultraspiracle (USP); wherein the DNA binding domain and the transactivation domain are from other than EcR, RXR, or USP; wherein the ligand binding domains from the first polypeptide and the second polypeptide are different and dimerize.

**[0127]** The present invention also relates to a gene expression modulation system according to the present invention further comprising c) a third gene expression cassette comprising: i) the response element to which the DNA-binding domain of the first polypeptide binds; ii) a promoter that is activated by the transactivation domain of the second polypeptide; and iii) the gene whose expression is to be modulated.

**[0128]** In a specific embodiment, the gene whose expression is to be modulated is a homologous gene with respect

to the host cell. In another specific embodiment, the gene whose expression is to be modulated is a heterologous gene with respect to the host cell.

**[0129]** In a specific embodiment, the ligand binding domain of the first polypeptide comprises an ecdysone receptor ligand binding domain.

**[0130]** In another specific embodiment, the ligand binding domain of the first polypeptide comprises a retinoid X receptor ligand binding domain.

**[0131]** In a specific embodiment, the ligand binding domain of the second polypeptide comprises an ecdysone receptor ligand binding domain.

**[0132]** In another specific embodiment, the ligand binding domain of the second polypeptide comprises a retinoid X receptor ligand binding domain.

**[0133]** In a preferred embodiment, the ligand binding domain of the first polypeptide comprises an ecdysone receptor ligand binding domain, and the ligand binding domain of the second polypeptide comprises a retinoid X receptor ligand binding domain.

**[0134]** In another preferred embodiment, the ligand binding domain of the first polypeptide is from a retinoid X receptor polypeptide, and the ligand binding domain of the second polypeptide is from an ecdysone receptor polypeptide.

**[0135]** Preferably, the ligand binding domain is an EcR or RXR related steroid/thyroid hormone nuclear receptor family member ligand binding domain, or analogs, combinations, or modifications thereof. More preferably, the LBD is from EcR or RXR. Even more preferably, the LBD is from a truncated EcR or RXR. A truncation mutation may be made by any method used in the art, including but not limited to restriction endonuclease digestion/deletion, PCR-mediated/oligonucleotide-directed deletion, chemical mutagenesis, UV strand breakage, and the like.

**[0136]** Preferably, the EcR is an insect EcR selected from the group consisting of a Lepidopteran EcR, a Dipteran EcR, an Arthropod EcR, a Homopteran EcR and a Hemipteran EcR. More preferably, the EcR for use is a spruce budworm *Choristoneura fumiferana* EcR ("CfEcR"), a *Tenebrio molitor* EcR ("TmEcR"), a *Manduca sexta* EcR ("MsEcR"), a *Heliothis virescens* EcR ("HvEcR"), a silk moth *Bombyx mori* EcR ("BmEcR"), a fruit fly *Drosophila melanogaster* EcR ("DmEcR"), a mosquito *Aedes aegypti* EcR ("AaEcR"), a blowfly *Lucilia capitata* EcR ("LcEcR"), a Mediterranean fruit fly *Ceratitis capitata* EcR ("CcEcR"), a locust *Locusta migratoria* EcR ("LmEcR"), an aphid *Myzus persicae* EcR ("MpEcR"), a fiddler crab *Uca pugilator* EcR ("UpEcR"), or an ixodid tick *Amblyomma americanum* EcR ("AmaEcR"). Even more preferably, the LBD is from spruce budworm (*Choristoneura fumiferana*) EcR ("CfEcR") or fruit fly *Drosophila melanogaster* EcR ("DmEcR").

**[0137]** Preferably, the LBD is from a truncated insect EcR. The insect EcR polypeptide truncation comprises a deletion of at least 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, or 265 amino acids. More preferably, the insect EcR polypeptide truncation comprises a deletion of at least a partial polypeptide domain. Even more preferably, the insect EcR polypeptide truncation comprises a deletion of at least an entire polypeptide domain. In a specific embodiment, the insect EcR polypeptide truncation comprises a deletion of at least an A/B-domain deletion, a C-domain deletion, a D-domain deletion, an E-domain deletion, an F-domain deletion, an A/B/C-domains deletion, an AB/1/2-C-domains deletion, an A/B/C/D-domains deletion, an AB/C/D/F-domains deletion, an A/B/F-domains, and an A/B/C/F-domains deletion. A combination of several complete and/or partial domain deletions may also be performed.

**[0138]** In a preferred embodiment, the ecdysone receptor ligand binding domain is encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, and SEQ ID NO: 10.

**[0139]** In another preferred embodiment, the ecdysone receptor ligand binding domain comprises a polypeptide sequence selected from the group consisting of SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, and SEQ ID NO: 20.

**[0140]** Preferably, the RXR polypeptide is a mouse *Mus musculus* RXR ("MmRXR") or a human *Homo sapiens* RXR ("HsRXR"). The RXR polypeptide may be an RXR $_{\alpha}$ , RXR $_{\beta}$ , or RXR $_{\gamma}$  isoform.

**[0141]** Preferably, the LBD is from a truncated RXR. The RXR polypeptide truncation comprises a deletion of at least 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, or 265 amino acids. More preferably, the RXR polypeptide truncation comprises a deletion of at least a partial polypeptide domain. Even more preferably, the RXR polypeptide truncation comprises a deletion of at least an entire polypeptide domain. In a specific embodiment, the RXR polypeptide truncation comprises a deletion of at least an A/B-domain deletion, a C-domain deletion, a D-domain deletion, an E-domain deletion, an F-domain deletion, an A/B/C-domains deletion, an AB/1/2-C-domains deletion, an A/B/C/D-domains deletion, an AB/C/D/F-domains deletion, an A/B/F-domains, and an A/B/C/F-domains deletion. A combination of several complete and/or partial domain deletions may also be performed.

**[0142]** In a preferred embodiment, the retinoid X receptor ligand binding domain is encoded by a polynucleotide

comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, and SEQ ID NO: 30.

[0143] In another preferred embodiment, the retinoid X receptor ligand binding domain comprises a polypeptide sequence selected from the group consisting of SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, and SEQ ID NO: 40.

[0144] For purposes of this invention EcR and RXR also include synthetic and chimeric EcR and RXR and their homologs.

[0145] The DNA binding domain can be any DNA binding domain with a known response element, including synthetic and chimeric DNA binding domains, or analogs, combinations, or modifications thereof. Preferably, the DBD is a GAL4 DBD, a LexA DBD, a transcription factor DBD, a steroid/thyroid hormone nuclear receptor superfamily member DBD, a bacterial LacZ DBD, or a yeast po DBD. More preferably, the DBD is a GAL4 DBD [SEQ ID NO: 41 (polynucleotide) or SEQ ID NO: 42 (polypeptide)] or a LexA DBD [(SEQ ID NO: 43 (polynucleotide) or SEQ ID NO: 44 (polypeptide)].

[0146] The transactivation domain (abbreviated "AD" or "TA") may be any steroid/thyroid hormone nuclear receptor AD, synthetic or chimeric AD, polyglutamine AD, basic or acidic amino acid AD, a VP16 AD, a GAL4 AD, an NF- $\kappa$ B AD, a BP64 AD, or an analog, combination, or modification thereof. Preferably, the AD is a synthetic or chimeric AD, or is obtained from a VP16, GAL4, or NF- $\kappa$ B. Most preferably, the AD is a VP16 AD [SEQ ID NO: 45 (polynucleotide) or SEQ ID NO: 46 (polypeptide)].

[0147] The response element ("RE") may be any response element with a known DNA binding domain, or an analog, combination, or modification thereof. Preferably, the RE is an RE from GAL4 ("GAL4RE"), LexA, a steroid/thyroid hormone nuclear receptor RE, or a synthetic RE that recognizes a synthetic DNA binding domain. More preferably, the RE is a GAL4RE comprising a polynucleotide sequence of SEQ ID NO: 47 or a LexA 8X operon comprising a polynucleotide sequence of SEQ ID NO: 48. Preferably, the first hybrid protein is substantially free of a transactivation domain and the second hybrid protein is substantially free of a DNA binding domain. For purposes of this invention, "substantially free" means that the protein in question does not contain a sufficient sequence of the domain in question to provide activation or binding activity.

[0148] The ligands for use in the present invention as described below, when combined with the ligand binding domain of an EcR, USP, RXR, or another polypeptide which in turn are bound to the response element linked to a gene, provide the means for external temporal regulation of expression of the gene. The binding mechanism or the order in which the various components of this invention bind to each other, that is, ligand to receptor, first polypeptide to response element, second polypeptide to promoter, etc., is not critical. Binding of the ligand to the ligand binding domains of an EcR, USP, RXR, or another protein, enables expression or suppression of the gene. This mechanism does not exclude the potential for ligand binding to EcR, USP, or RXR, and the resulting formation of active homodimer complexes (e.g. EcR+EcR or USP+USP). Preferably, one or more of the receptor domains can be varied producing a chimeric gene switch. Typically, one or more of the three domains, DBD, LBD, and transactivation domain, may be chosen from a source different than the source of the other domains so that the chimeric genes and the resulting hybrid proteins are optimized in the chosen host cell or organism for transactivating activity, complementary binding of the ligand, and recognition of a specific response element. In addition, the response element itself can be modified or substituted with response elements for other DNA binding protein domains such as the GAL-4 protein from yeast (see Sadowski, et al. (1988) Nature, 335: 563-564) or LexA protein from *E. coli* (see Brent and Ptashne (1985), Cell, 43:729-736), or synthetic response elements specific for targeted interactions with proteins designed, modified, and selected for such specific interactions (see, for example, Kim, et al. (1997), *Proc. Natl. Acad. Sci., USA*, 94:3616-3620) to accommodate chimeric receptors. Another advantage of chimeric systems is that they allow choice of a promoter used to drive the gene expression according to a desired end result. Such double control can be particularly important in areas of gene therapy, especially when cytotoxic proteins are produced, because both the timing of expression as well as the cells wherein expression occurs can be controlled. When genes, operatively linked to a suitable promoter, are introduced into the cells of the subject, expression of the exogenous genes is controlled by the presence of the system of this invention. Promoters may be constitutively or inducibly regulated or may be tissue-specific (that is, expressed only in a particular type of cells) or specific to certain developmental stages of the organism.

## GENE EXPRESSION CASSETTES OF THE INVENTION

[0149] The novel ecdysone receptor-based inducible gene expression system of the invention comprises a novel gene expression cassette that is capable of being expressed in a host cell, wherein the gene expression cassette comprises a polynucleotide encoding a hybrid polypeptide. Thus, Applicants' invention also provides novel gene expression cassettes for use in the gene expression system of the invention.

[0150] Specifically, the present invention provides a gene expression cassette comprising a polynucleotide encoding a hybrid polypeptide. The hybrid polypeptide comprises either 1) a DNA-binding domain that recognizes a response element and a ligand binding domain of a nuclear receptor or 2) a transactivation domain and a ligand binding domain

of a nuclear receptor, wherein the transactivation domain is from a nuclear receptor other than an EcR, an RXR, or a USP.

**[0151]** In a specific embodiment, the gene expression cassette encodes a hybrid polypeptide comprising a DNA-binding domain that recognizes a response element and an ecdysone receptor ligand binding domain, wherein the DNA binding domain is from a nuclear receptor other than an ecdysone receptor.

**[0152]** In another specific embodiment, the gene expression cassette encodes a hybrid polypeptide comprising a DNA binding domain that recognizes a response element and a retinoid X receptor ligand binding domain, wherein the DNA binding domain is from a nuclear receptor other than a retinoid X receptor.

**[0153]** The DNA binding domain can be any DNA binding domain with a known response element, including synthetic and chimeric DNA binding domains, or analogs, combinations, or modifications thereof. Preferably, the DBD is a GAL4 DBD, a LexA DBD, a transcription factor DBD, a steroid/thyroid hormone nuclear receptor superfamily member DBD, a bacterial LacZ DBD, or a yeast put DBD. More preferably, the DBD is a GAL4 DBD [SEQ ID NO: 41 (polynucleotide) or SEQ ID NO: 42 (polypeptide)] or a LexA DBD [SEQ ID NO: 43 (polynucleotide) or SEQ ID NO: 44 (polypeptide)].

**[0154]** In another specific embodiment, the gene expression cassette encodes a hybrid polypeptide comprising a transactivation domain and an ecdysone receptor ligand binding domain, wherein the transactivation domain is from a nuclear receptor other than an ecdysone receptor.

**[0155]** In another specific embodiment, the gene expression cassette encodes a hybrid polypeptide comprising a transactivation domain and a retinoid X receptor ligand binding domain, wherein the transactivation domain is from a nuclear receptor other than a retinoid X receptor.

**[0156]** The transactivation domain (abbreviated "AD" or "TA") may be any steroid/thyroid hormone nuclear receptor AD, synthetic or chimeric AD, polyglutamine AD, basic or acidic amino acid AD, a VP16 AD, a GAL4 AD, an NF-κB AD, a BP64 AD, or an analog, combination, or modification thereof. Preferably, the AD is a synthetic or chimeric AD, or is obtained from a VP16, GAL4, or NF-κB. Most preferably, the AD is a VP16 AD [SEQ ID NO: 45 (polynucleotide) or SEQ ID NO: 46 (polypeptide)].

**[0157]** Preferably, the ligand binding domain is an EcR or RXR related steroid/thyroid hormone nuclear receptor family member ligand binding domain, or analogs, combinations, or modifications thereof. More preferably, the LBD is from an EcR or RXR. Even more preferably, the LBD is from a truncated EcR or RXR.

**[0158]** Preferably, the EcR is an insect EcR selected from the group consisting of a Lepidopteran EcR, a Dipteran EcR, an Arthropod EcR, a Homopteran EcR and a Hemipteran EcR. More preferably, the EcR for use is a spruce budworm *Choristoneura fumiferana* EcR ("CfEcR"), a *Tenebrio molitor* EcR ("TmEcR"), a *Manduca sexta* EcR ("MsEcR"), a *Heliothis virescens* EcR ("HvEcR"), a silk moth *Bombyx mori* EcR ("BmEcR"), a fruit fly *Drosophila melanogaster* EcR ("DmEcR"), a mosquito *Aedes aegypti* EcR ("AaEcR"), a blowfly *Lucilia capitata* EcR ("LcEcR"), a Mediterranean fruit fly *Ceratitis capitata* EcR ("CcEcR"), a locust *Locusta migratoria* EcR ("LmEcR"), an aphid *Myzus persicae* EcR ("MpEcR"), a fiddler crab *Uca pugnator* EcR ("UpEcR"), or an iodoxid tick *Amblyomma americanum* EcR ("AmaEcR"). Even more preferably, the LBD is from spruce budworm (*Choristoneura fumiferana*) EcR ("CfEcR") or fruit fly *Drosophila melanogaster* EcR ("DmEcR").

**[0159]** Preferably, the LBD is from a truncated insect EcR. The insect EcR polypeptide truncation comprises a deletion of at least 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 195, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, or 265 amino acids. More preferably, the insect EcR polypeptide truncation comprises a deletion of at least a partial polypeptide domain. Even more preferably, the insect EcR polypeptide truncation comprises a deletion of at least an entire polypeptide domain. In a specific embodiment, the insect EcR polypeptide truncation comprises a deletion of at least an A/B-domain deletion, a C-domain deletion, a D-domain deletion, an E-domain deletion, an F-domain deletion, an A/B/C-domains deletion, an AB/1/2-C-domains deletion, an A/B/C/D-domains deletion, an A/B/C/D/F-domains deletion, an A/B/F-domains, and an A/B/C/F-domains deletion. A combination of several complete and/or partial domain deletions may also be performed.

**[0160]** In a preferred embodiment, the ecdysone receptor ligand binding domain is encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, and SEQ ID NO: 10.

**[0161]** In another preferred embodiment, the ecdysone receptor ligand binding domain comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, and SEQ ID NO: 20.

**[0162]** Preferably, the RXR polypeptide is a mouse *Mus musculus* RXR ("MmRXR") or a human *Homo sapiens* RXR ("HsRXR"). The RXR polypeptide may be an RXR<sub>α</sub>, RXR<sub>β</sub>, or RXR<sub>γ</sub> isoform.

**[0163]** Preferably, the LBD is from a truncated RXR. The RXR polypeptide truncation comprises a deletion of at least 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, or 265 amino acids. More preferably, the RXR polypeptide truncation comprises a deletion of at least a partial polypeptide domain. Even more preferably, the RXR polypeptide truncation comprises a deletion of at least an entire

polypeptide domain. In a specific embodiment, the RXR polypeptide truncation comprises a deletion of at least an A/B-domain deletion, a C-domain deletion, a D-domain deletion, an E-domain deletion, an F-domain deletion, an A/B/C-domains deletion, an A/B1/2-C-domains deletion, an A/B/C/D-domains deletion, an A/B/C/D/P-domains deletion, an A/B/F-domains, and an A/B/C/F-domains deletion. A combination of several complete and/or partial domain deletions may also be performed.

**[0164]** In a preferred embodiment, the retinoid X receptor ligand binding domain is encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, and SEQ ID NO: 30.

**[0165]** In another preferred embodiment, the retinoid X receptor ligand binding domain comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, and SEQ ID NO: 40.

**[0166]** In a preferred embodiment, the gene expression cassette encodes a hybrid polypeptide comprising a DNA-binding domain encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of a GAL4 DBD (SEQ ID NO: 41) or a LexA DBD (SEQ ID NO: 43) and an ecdysone receptor ligand binding domain encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, and SEQ ID NO: 10.

**[0167]** In another preferred embodiment, the gene expression cassette encodes a hybrid polypeptide comprising a DNA-binding domain comprising a polypeptide sequence selected from the group consisting of a GAL4 DBD (SEQ ID NO: 42) or a LexA DBD (SEQ ID NO: 44) and an ecdysone receptor ligand binding domain comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, and SEQ ID NO: 20.

**[0168]** In another preferred embodiment, the gene expression cassette encodes a hybrid polypeptide comprising a DNA-binding domain encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of a GAL4 DBD (SEQ ID NO: 41) or a LexA DBD (SEQ ID NO: 43) and a retinoid X receptor ligand binding domain encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, and SEQ ID NO: 30.

**[0169]** In another preferred embodiment, the gene expression cassette encodes a hybrid polypeptide comprising a DNA-binding domain comprising a polypeptide sequence selected from the group consisting of a GAL4 DBD (SEQ ID NO: 42) or a LexA DBD (SEQ ID NO: 44) and a retinoid X receptor ligand binding domain comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, and SEQ ID NO: 40.

**[0170]** In another preferred embodiment, the gene expression cassette encodes a hybrid polypeptide comprising a transactivation domain encoded by a polynucleotide comprising a nucleic acid sequence of SEQ ID NO: 45 and an ecdysone receptor ligand binding domain encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, and SEQ ID NO: 10.

**[0171]** In another preferred embodiment, the gene expression cassette encodes a hybrid polypeptide comprising a transactivation domain comprising a polypeptide sequence of SEQ ID NO: 46 and an ecdysone receptor ligand binding domain comprising a polypeptide sequence selected from the group consisting of SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, and SEQ ID NO: 20.

**[0172]** In another preferred embodiment, the gene expression cassette encodes a hybrid polypeptide comprising a transactivation domain encoded by a polynucleotide comprising a nucleic acid sequence of SEQ ID NO: 45 and a retinoid X receptor ligand binding domain encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, and SEQ ID NO: 30.

**[0173]** In another preferred embodiment, the gene expression cassette encodes a hybrid polypeptide comprising a transactivation domain comprising a polypeptide sequence of SEQ ID NO: 46 and a retinoid X receptor ligand binding domain comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, and SEQ ID NO: 40.

**[0174]** For purposes of this invention EcR and RXR also include synthetic and chimeric EcR and RXR and their homologs.

## POLYNUCLEOTIDES OF THE INVENTION

**[0175]** The novel ecdysone receptor-based inducible gene expression system of the invention comprises a gene expression cassette comprising a polynucleotide that encodes a truncated EcR or RXR polypeptide comprising a truncation mutation and is useful in methods of modulating the expression of a gene within a host cell.

**[0176]** Thus, the present invention also relates to a polynucleotide that encodes an EcR or RXR polypeptide comprising a truncation mutation. Specifically, the present invention relates to an isolated polynucleotide encoding an EcR or RXR polypeptide comprising a truncation mutation that affects ligand binding activity or ligand sensitivity.

**[0177]** Preferably, the truncation mutation results in a polynucleotide that encodes a truncated EcR polypeptide or a truncated RXR polypeptide comprising a deletion of at least 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, or 265 amino acids. More preferably, the EcR or RXR polypeptide truncation comprises a deletion of at least a partial polypeptide domain. Even more preferably, the EcR or RXR polypeptide truncation comprises a deletion of at least an entire polypeptide domain. In a specific embodiment, the EcR or RXR polypeptide truncation comprises a deletion of at least an A/B-domain deletion, a C-domain deletion, a D-domain deletion, an E-domain deletion, an F-domain deletion, an A/B/C-domains deletion, an A/B/1/2-C-domains deletion, an A/B/C/D-domains deletion, an A/B/C/D/F-domains deletion, an A/B/F-domains, and an A/B/C/F-domains deletion. A combination of several complete and/or partial domain deletions may also be performed.

**[0178]** In a specific embodiment, the EcR polynucleotide according to the invention comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, and SEQ ID NO: 10. In a specific embodiment, the polynucleotide according to the invention encodes a ecdysone receptor polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 11 (CfEcR-CDEF), SEQ ID NO: 12 (CfEcR-1/2CDEF, which comprises the last 33 carboxy-terminal amino acids of C domain), SEQ ID NO: 13 (CfEcR-DEF), SEQ ID NO: 14 (CfEcR-EF), SEQ ID NO: 15 (CfEcR-DE), SEQ ID NO: 16 (DmEcR-CDEF), SEQ ID NO: 17 (DmEcR-1/2CDEF), SEQ ID NO: 18 (DmEcR-DEF), SEQ ID NO: 19 (DmEcR-EF), and SEQ ID NO: 20 (DmEcR-DE).

**[0179]** In another specific embodiment, the RXR polynucleotide according to the invention comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, and SEQ ID NO: 30. In another specific embodiment, the polynucleotide according to the invention encodes a truncated RXR polypeptide comprising an amino acid sequence consisting of SEQ ID NO: 31 (MmRXR-CDEF), SEQ ID NO: 32 (MmRXR-DEF), SEQ ID NO: 33 (MmRXR-EF), SEQ ID NO: 34 (MmRXR-truncatedEF), SEQ ID NO: 35 (MmRXR-E), SEQ ID NO: 36 (HsRXR-CDEF), SEQ ID NO: 37 (HsRXR-DEF), SEQ ID NO: 38 (HsRXR-EF), SEQ ID NO: 39 (HsRXR-truncated EF), and SEQ ID NO: 40 (HsRXR-E).

**[0180]** In particular, the present invention relates to an isolated polynucleotide encoding an EcR or RXR polypeptide comprising a truncation mutation, wherein the mutation reduces ligand binding activity or ligand sensitivity of the EcR or RXR polypeptide. In a specific embodiment, the present invention relates to an isolated polynucleotide encoding an EcR or RXR polypeptide comprising a truncation mutation that reduces steroid binding activity or steroid sensitivity of the EcR or RXR polypeptide. In a preferred embodiment, the present invention relates to an isolated polynucleotide encoding an EcR polypeptide comprising a truncation mutation that reduces steroid binding activity or steroid sensitivity of the EcR polypeptide, wherein the polynucleotide comprises a nucleic acid sequence of SEQ ID NO: 3 (CfEcR-DEF), SEQ ID NO: 4 (CfEcR-EF), SEQ ID NO: 8 (DmEcR-DEF), or SEQ ID NO: 9 (DmEcR-EF). In another specific embodiment, the present invention relates to an isolated polynucleotide encoding an EcR or RXR polypeptide comprising a truncation mutation that reduces non-steroid binding activity or non-steroid sensitivity of the EcR or RXR polypeptide. In a preferred embodiment, the present invention relates to an isolated polynucleotide encoding an EcR polypeptide comprising a truncation mutation that reduces non-steroid binding activity or non-steroid sensitivity of the EcR polypeptide, wherein the polynucleotide comprises a nucleic acid sequence of SEQ ID NO: 4 (CfEcR-EF) or SEQ ID NO: 9 (DmEcR-EF).

**[0181]** The present invention also relates to an isolated polynucleotide encoding an EcR or RXR polypeptide comprising a truncation mutation, wherein the mutation enhances ligand binding activity or ligand sensitivity of the EcR or RXR polypeptide. In a specific embodiment, the present invention relates to an isolated polynucleotide encoding an EcR or RXR polypeptide comprising a truncation mutation that enhances steroid binding activity or steroid sensitivity of the EcR or RXR polypeptide. In another specific embodiment, the present invention relates to an isolated polynucleotide encoding an EcR or RXR polypeptide comprising a truncation mutation that enhances non-steroid binding activity or non-steroid sensitivity of the EcR or RXR polypeptide. In a preferred embodiment, the present invention relates to an isolated polynucleotide encoding an EcR polypeptide comprising a truncation mutation that enhances non-steroid binding activity or non-steroid sensitivity of the EcR polypeptide, wherein the polynucleotide comprises a nucleic acid sequence of SEQ ID NO: 3 (CfEcR-DEF) or SEQ ID NO: 8 (DmEcR-DEF).

**[0182]** The present invention also relates to an isolated polynucleotide encoding a retinoid X receptor polypeptide

comprising a truncation mutation that increases ligand sensitivity of a heterodimer comprising the mutated retinoid X receptor polypeptide and a dimerization partner. Preferably, the isolated polynucleotide encoding a retinoid X receptor polypeptide comprising a truncation mutation that increases ligand sensitivity of a heterodimer comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO: 23 (MmRXR-EF), SEQ ID NO: 24 (114mRXR-truncatedEF), SEQ ID NO: 28 (HsRXR-EF), or SEQ ID NO: 29 (HsRXR-truncated EF). In a specific embodiment, the dimerization partner is an ecdysone receptor polypeptide. Preferably, the dimerization partner is a truncated EcR polypeptide. More preferably, the dimerization partner is an EcR polypeptide in which domains A/B/C have been deleted. Even more preferably, the dimerization partner is an EcR polypeptide comprising an amino acid sequence of SEQ ID NO: 13 (CfEcR-DEF) or SEQ ID NO: 18 (DmEcR-DEF).

## POLYPEPTIDES OF THE INVENTION

**[0183]** The novel ecdysone receptor-based inducible gene expression system of the invention comprises a polynucleotide that encodes a truncated EcR or RXR polypeptide and is useful in methods of modulating the expression of a gene within a host cell. Thus, the present invention also relates to an isolated truncated EcR or RXR polypeptide encoded by a polynucleotide or a gene expression cassette according to the invention. Specifically, the present invention relates to an isolated truncated EcR or RXR polypeptide comprising a truncation mutation that affects ligand binding activity or ligand sensitivity encoded by a polynucleotide according to the invention.

**[0184]** The present invention also relates to an isolated truncated EcR or RXR polypeptide comprising a truncation mutation. Specifically, the present invention relates to an isolated EcR or RXR polypeptide comprising a truncation mutation that affects ligand binding activity or ligand sensitivity.

**[0185]** Preferably, the truncation mutation results in a truncated EcR polypeptide or a truncated RXR polypeptide comprising a deletion of at least 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, or 265 amino acids. More preferably, the EcR or RXR polypeptide truncation comprises a deletion of at least a partial polypeptide domain. Even more preferably, the EcR or RXR polypeptide truncation comprises a deletion of at least an entire polypeptide domain. In a specific embodiment, the EcR or RXR polypeptide truncation comprises a deletion of at least an A/B-domain deletion, a C-domain deletion, a D-domain deletion, an E-domain deletion, an F-domain deletion, an A/B/C-domains deletion, an A/B/1/2-C-domains deletion, an AB/C/D-domains deletion, an AB/C/D/F-domains deletion, an A/B/F-domains, and an A/B/C/F-domains deletion. A combination of several complete and/or partial domain deletions may also be performed.

**[0186]** In a preferred embodiment, the isolated truncated ecdysone receptor polypeptide is encoded by a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO: 1 (CfEcR-CDEF), SEQ ID NO: 2 (CfEcR-1/2CDEF), SEQ ID NO: 3 (CfEcR-DEF), SEQ ID NO: 4 (CfEcR-EF), SEQ ID NO: 5 (CfEcR-DE), SEQ ID NO: 6 (DmEcR-CDEF), SEQ ID NO: 7 (DmEcR-1/2CDEF), SEQ ID NO: 8 (DmEcR-DEF), SEQ ID NO: 9 (DmEcR-EF), and SEQ ID NO: 10 (DmEcR-DE). In another preferred embodiment, the isolated truncated ecdysone receptor polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 11 (CfEcR-CDEF), SEQ ID NO: 12 (CfEcR-1/2CDEF), SEQ ID NO: 13 (CfEcR-DEF), SEQ ID NO: 14 (CfEcR-EF), SEQ ID NO: 15 (CfEcR-DE), SEQ ID NO: 16 (DmEcR-CDEF), SEQ ID NO: 17 (DmEcR-1/2CDEF), SEQ ID NO: 18 (DmEcR-DEF), SEQ ID NO: 19 (DmEcR-EF), and SEQ ID NO: 20 (DmEcR-DE).

**[0187]** In a preferred embodiment, the isolated truncated RXR polypeptide is encoded by a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO: 21 (MmRXR-CDEF), SEQ ID NO: 22 (MmRXR-DEF), SEQ ID NO: 23 (MmRXR-EF), SEQ ID NO: 24 (MmRXR-truncatedEF), SEQ ID NO: 25 (MmRXR-E), SEQ ID NO: 26 (HsRXR-CDEF), SEQ ID NO: 27 (HsRXR-DEF), SEQ ID NO: 28 (HsRXR-EF), SEQ ID NO: 29 (HsRXR-truncatedEF) and SEQ ID NO: 30 (HsRXR-E). In another preferred embodiment, the isolated truncated RXR polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 31 (MmRXR-CDEF), SEQ ID NO: 32 (MmRXR-DEF), SEQ ID NO: 33 (MmRXR-EF), SEQ ID NO: 34 (MmRXR-truncatedEF), SEQ ID NO: 35 (MmRXR-E), SEQ ID NO: 36 (HsRXR-CDEF), SEQ ID NO: 37 (HsRXR-DEF), SEQ ID NO: 38 (HsRXR-EF), SEQ ID NO: 39 (HsRXR-truncatedEF), and SEQ ID NO: 40 (HsRXR-E).

**[0188]** The present invention relates to an isolated EcR or RXR polypeptide comprising a truncation mutation that reduces ligand binding activity or ligand sensitivity of the EcR or RXR polypeptide, wherein the polypeptide is encoded by a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO: 1 (CfEcR-CDEF), SEQ ID NO: 2 (CfEcR-1/2CDEF), SEQ ID NO: 3 (CfEcR-DEF), SEQ ID NO: 4 (CfEcR-EF), SEQ ID NO: 5 (CfEcR-DE), SEQ ID NO: 6 (DmEcR-CDEF), SEQ ID NO: 7 (DmEcR-1/2CDEF), SEQ ID NO: 8 (DmEcR-DEF), SEQ ID NO: 9 (DmEcR-EF), SEQ ID NO: 10 (DmEcR-DE), SEQ ID NO: 21 (MmRXR-DEF), SEQ ID NO: 22 (MmRXR-DEF), SEQ ID NO: 23 (MmRXR-EF), SEQ ID NO: 24 (MmRXR-truncatedEF), SEQ ID NO: 25 (MmRXR-E), SEQ ID NO: 26 (HsRXR-CDEF), SEQ ID NO: 27 (HsRXR-DEF), SEQ ID NO: 28 (HsRXR-EF), SEQ ID NO: 29 (HsRXR-truncatedEF), and SEQ ID NO: 30 (HsRXR-E).

**[0189]** Thus, the present invention relates to an isolated truncated EcR or RXR polypeptide comprising a truncation mutation that reduces ligand binding activity or ligand sensitivity of the EcR or RXR polypeptide, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 11 (CfEcR-CDEF), SEQ ID NO: 12 (CfEcR-1/2CDEF), SEQ ID NO: 13 (CfEcR-DEF), SEQ ID NO: 14 (CfEcR-EF), SEQ ID NO: 15 (CfEcR-DE), SEQ ID NO: 16 (DmEcR-CDEF), SEQ ID NO: 17 (DmEcR-1/2CDEF), SEQ ID NO: 18 (DmEcR-DEF), SEQ ID NO: 19 (DmEcR-EF), SEQ ID NO: 20 (DmEcR-DE), SEQ ID NO: 31 (MmRXR-CDEF), SEQ ID NO: 32 (MmRXR-DEF), SEQ ID NO: 33 (MmRXR-EF), SEQ ID NO: 34 (MmRXR-truncatedEF), SEQ ID NO: 35 (MmRXR-E), SEQ ID NO: 36 (HsRXR-CDEF), SEQ ID NO: 37 (HsRXR-DEF), SEQ ID NO: 38 (HsRXR-EF), SEQ ID NO: 39 (HsRXR-truncatedEF), and SEQ ID NO: 40 (HsRXR-E).

**[0190]** In a specific embodiment, the present invention relates to an isolated EcR or RXR polypeptide comprising a truncation mutation that reduces steroid binding activity or steroid sensitivity of the EcR or RXR polypeptide. In a preferred embodiment, the present invention relates to an isolated EcR polypeptide comprising a truncation mutation that reduces steroid binding activity or steroid sensitivity of the EcR polypeptide, wherein the EcR polypeptide is encoded by a polynucleotide comprising a nucleic acid sequence of SEQ ID NO: 3 (CfEcR-DEF), SEQ ID NO: 4 (CfEcR-EF), SEQ ID NO: 8 (DmEcR-DEF), or SEQ ID NO: 9 (DmEcR-EF). Accordingly, the present invention also relates to an isolated truncated EcR or RXR polypeptide comprising a truncation mutation that reduces steroid binding activity or steroid sensitivity of the EcR or RXR polypeptide. In a preferred embodiment, the present invention relates to an isolated EcR polypeptide comprising a truncation mutation that reduces steroid binding activity or steroid sensitivity of the EcR polypeptide, wherein the EcR polypeptide comprises an amino acid sequence of SEQ ID NO: 13 (CfEcR-DEF), SEQ ID NO: 14 (CfEcR-EF), SEQ ID NO: 18 (DmEcR-DEF), or SEQ ID NO: 19 (DmEcR-EF).

**[0191]** In another specific embodiment, the present invention relates to an isolated EcR or RXR polypeptide comprising a truncation mutation that reduces non-steroid binding activity or non-steroid sensitivity of the EcR or RXR polypeptide. In a preferred embodiment, the present invention relates to an isolated EcR polypeptide comprising a truncation mutation that reduces non-steroid binding activity or non-steroid sensitivity of the EcR polypeptide, wherein the EcR polypeptide is encoded by a polynucleotide comprising a nucleic acid sequence of SEQ ID NO: 4 (CfEcR-EF) or SEQ ID NO: 9 (DmEcR-EF). Accordingly, the present invention also relates to an isolated truncated EcR or RXR polypeptide comprising a truncation mutation that reduces non-steroid binding activity or steroid sensitivity of the EcR or RXR polypeptide. In a preferred embodiment, the present invention relates to an isolated EcR polypeptide comprising a truncation mutation that reduces non-steroid binding activity or non-steroid sensitivity of the EcR polypeptide, wherein the EcR polypeptide comprises an amino acid sequence of SEQ ID NO: 14 (CfEcR-EF) or SEQ ID NO: 19 (DmEcR-EF).

**[0192]** In particular, the present invention relates to an isolated EcR or RXR polypeptide comprising a truncation mutation that enhances ligand binding activity or ligand sensitivity of the EcR or RXR polypeptide, wherein the polypeptide is encoded by a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO: 1 (CfEcR-CDEF), SEQ ID NO: 2 (CfEcR-1/2CDEF), SEQ ID NO: 3 (CfEcR-DEF), SEQ ID NO: 4 (CfEcR-EF), SEQ ID NO: 5 (CfEcR-DE), SEQ ID NO: 6 (DmEcR-CDEF), SEQ ID NO: 7 (DmEcR-1/2CDEF), SEQ ID NO: 8 (DmEcR-DEF), SEQ ID NO: 9 (DmEcR-EF), SEQ ID NO: 10 (DmEcR-DE), SEQ ID NO: 21 (MmRXR-CDEF), SEQ ID NO: 22 (MmRXR-DEF), SEQ ID NO: 23 (MmRXR-EF), SEQ ID NO: 24 (MmRXR-truncatedEF), SEQ ID NO: 25 (MmRXR-E), SEQ ID NO: 26 (HsRXR-CDEF), SEQ ID NO: 27 (HsRXR-DEF), SEQ ID NO: 28 (HsRXR-EF), SEQ ID NO: 29 (HsRXR-truncatedEF), and SEQ ID NO: 30 (HsRXR-E).

**[0193]** The present invention relates to an isolated EcR or RXR polypeptide comprising a truncation mutation that enhances ligand binding activity or ligand sensitivity of the EcR or RXR polypeptide, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 11 (CfEcR-CDEF), SEQ ID NO: 12 (CfEcR-1/2CDEF), SEQ ID NO: 13 (CfEcR-DEF), SEQ ID NO: 14 (CfEcR-EF), SEQ ID NO: 15 (CfEcR-DE), SEQ ID NO: 16 (DmEcR-CDEF), SEQ ID NO: 17 (DmEcR-1/2CDEF), SEQ ID NO: 18 (DmEcR-DEF), SEQ ID NO: 19 (DmEcR-EF), SEQ ID NO: 20 (DmEcR-DE), SEQ ID NO: 31 (MmRXR-CDEF), SEQ ID NO: 32 (MmRXR-DEF), SEQ ID NO: 33 (MmRXR-EF), SEQ ID NO: 34 (MmRXR-truncatedEF), SEQ ID NO: 35 (MmRXR-E), SEQ ID NO: 36 (HsRXR-CDEF), SEQ ID NO: 37 (HsRXR-DEF), SEQ ID NO: 39 (HsRXR-EF), SEQ ID NO: 39 (HsRXR-truncatedEF), and SEQ ID NO: 40 (HsRXR-E).

**[0194]** The present invention relates to an isolated EcR or RXR polypeptide comprising a truncation mutation that enhances ligand binding activity or ligand sensitivity of the EcR or RXR polypeptide. In a specific embodiment, the present invention relates to an isolated EcR or RXR polypeptide comprising a truncation mutation that enhances steroid binding activity or steroid sensitivity of the EcR or RXR polypeptide. Accordingly, the present invention also relates to an isolated EcR or RXR polypeptide comprising a truncation mutation that enhances steroid binding activity or steroid sensitivity of the EcR or RXR polypeptide.

**[0195]** In another specific embodiment, the present invention relates to an isolated EcR or RXR polypeptide comprising a truncation mutation that enhances non-steroid binding activity or non-steroid sensitivity of the EcR or RXR polypeptide. In a preferred embodiment, the present invention relates to an isolated EcR polypeptide comprising a truncation mutation that enhances non-steroid binding activity or non-steroid sensitivity of the EcR polypeptide, wherein the EcR polypeptide



is encoded by a polynucleotide comprising a nucleic acid sequence of SEQ ID NO: 3 (CfEcR-DEF) or SEQ ID NO: 8 (DmEcR-DEF). Accordingly, the present invention also relates to an isolated EcR or RXR polypeptide comprising a truncation mutation that enhances non-steroid binding activity or steroid sensitivity of the EcR or RXR polypeptide. In a preferred embodiment, the present invention relates to an isolated EcR polypeptide comprising a truncation mutation that enhances non-steroid binding activity or non-steroid sensitivity of the EcR polypeptide, wherein the EcR polynucleotide comprises an amino acid sequence of SEQ ID NO: 13 (CfEcR-DEF) or SEQ ID NO: 18 (DmEcR-DEF).

**[0196]** The present invention also relates to an isolated retinoid X receptor polypeptide comprising a truncation mutation that increases ligand sensitivity of a heterodimer comprising the mutated retinoid X receptor polypeptide and a dimerization partner. Preferably, the isolated retinoid X receptor polypeptide comprising a truncation mutation that increases ligand sensitivity of a heterodimer is encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 23 (MmRXR-EF), SEQ ID NO: 24 (MmRXR-truncatedEF), SEQ ID NO: 28 (HsRXR-EF), or SEQ ID NO: 29 (HsRXR-truncatedEF). More preferably, the isolated polynucleotide encoding a retinoid X receptor polypeptide comprising a truncation mutation that increases ligand sensitivity of a heterodimer comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 33 (MmRXR-EF), SEQ ID NO: 34 (MmRXR-truncatedEF), SEQ ID NO: 38 (HsRXR-EF), or SEQ ID NO: 39 (HsRXR-truncatedEF).

**[0197]** In a specific embodiment, the dimerization partner is an ecdysone receptor polypeptide. Preferably, the dimerization partner is a truncated EcR polypeptide. More preferably, the dimerization partner is an EcR polypeptide in which domains A/B/C have been deleted. Even more preferably, the dimerization partner is an EcR polypeptide comprising an amino acid sequence of SEQ ID NO: 13 (CfEcR-DEF) or SEQ ID NO: 18 (DmEcR-DEF).

## METHOD OF MODULATING GENE EXPRESSION OF THE INVENTION

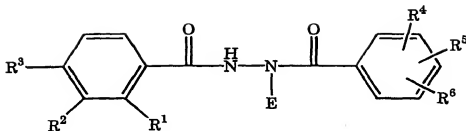
**[0198]** Applicants' invention also relates to methods of modulating gene expression in a host cell using a gene expression modulation system according to the invention. Specifically, Applicants' invention provides a method of modulating the expression of a gene in a host cell comprising the steps of: a) introducing into the host cell a gene expression modulation system according to the invention; and b) introducing into the host cell a ligand that independently combines with the ligand binding domains of the first polypeptide and the second polypeptide of the gene expression modulation system; wherein the gene to be expressed is a component of a gene expression cassette comprising: i) a response element comprising a domain to which the DNA binding domain of the first polypeptide binds; ii) a promoter that is activated by the transactivation domain of the second polypeptide; and iii) a gene whose expression is to be modulated, whereby a complex is formed comprising the ligand, the first polypeptide of the gene expression modulation system and the second polypeptide of the gene expression modulation system, and whereby the complex modulates expression of the gene in the host cell.

**[0199]** Genes of interest for expression in a host cell using Applicants' methods may be endogenous genes or heterologous genes. Nucleic acid or amino acid sequence information for a desired gene or protein can be located in one of many public access databases, for example, GENBANK, EMBL, Swiss-Prot, and PIR, or in many biology related journal publications. Thus, those skilled in the art have access to nucleic acid sequence information for virtually all known genes. Such information can then be used to construct the desired constructs for the insertion of the gene of interest within the gene expression cassettes used in Applicants' methods described herein.

**[0200]** Examples of genes of interest for expression in a host cell using Applicants' methods include, but are not limited to: antigens produced in plants as vaccines, enzymes like alpha-amylase, phytase, glucanase, and xylanase, genes for resistance against insects, nematodes, fungi, bacteria, viruses, and abiotic stresses, nutraceuticals, pharmaceuticals, vitamins, genes for modifying amino acid content, herbicide resistance, cold, drought, and heat tolerance, industrial products, oils, protein, carbohydrates, antioxidants, male sterile plants, flowers, fuels, other output traits, genes encoding therapeutically desirable polypeptides or products, such as genes that can provide, modulate, alleviate, correct and/or restore polypeptides important in treating a condition, a disease, a disorder, a dysfunction, a genetic defect, and the like.

**[0201]** Acceptable ligands are any that modulate expression of the gene when binding of the DNA binding domain of the two hybrid system to the response element in the presence of the ligand results in activation or suppression of expression of the genes. Preferred ligands include ponasterone, muristerone A, N,N'-diacylhydrazines such as those disclosed in U. S. Patents No. 6,013,836; 5,117,057; 5,530,028; and 5,378,726; dibenzoylalkyl cyanohydrazines such as those disclosed in European Application No. 461,809; N-alkyl-N,N'-diaryloxyhydrazines such as those disclosed in U. S. Patent No. 5,225,443; N-acyl-N-alkylcarboxylhydrazines such as those disclosed in European Application No. 234,994; N-aroxy-N-alkyl-N'-aroxyhydrazines such as those described in U. S. Patent No. 4,985,461; each of which is incorporated herein by reference and other similar materials including 3,5-di-tert-butyl-4-hydroxy-N-isobutyl-benzamide, 8-O-acetyl-harpagide, and the like.

**[0202]** Preferably, the ligand for use in Applicants' method of modulating expression of gene is a compound of the formula:



wherein:

- E is a (C<sub>4</sub>-C<sub>6</sub>)alkyl containing a tertiary carbon or a cyano(C<sub>3</sub>-C<sub>5</sub>)alkyl containing a tertiary carbon;  
 R<sup>1</sup> is H, Me, Et, i-Pr, F, formyl, CF<sub>3</sub>, CUF<sub>2</sub>, CHCl<sub>2</sub>, CH<sub>2</sub>F, CH<sub>2</sub>Cl, CH<sub>2</sub>OH, CH<sub>2</sub>OMe, CH<sub>2</sub>CN, CN, C<sup>o</sup>CH, 1-propynyl, 2-propynyl, vinyl, OH, OMe, OEt, cyclopropyl, CF<sub>2</sub>CF<sub>3</sub>, CH=CHCN, allyl, azido, SCN, or SCHF<sub>2</sub>;  
 R<sup>2</sup> is H, Me, Et, n-Pr, i-Pr, formyl, CF<sub>3</sub>, CHF<sub>2</sub>, CHCl<sub>2</sub>, CH<sub>2</sub>F, CH<sub>2</sub>Cl, CH<sub>2</sub>OH, CH<sub>2</sub>OMe, CH<sub>2</sub>CN, CN, C<sup>o</sup>CH, 1-propynyl, 2-propynyl, vinyl, Ac, F, Cl, OH, OMe, OEt, O-n-Pr, OAc, NMe<sub>2</sub>, NEt<sub>2</sub>, SMe, SEt, SOCF<sub>3</sub>, OCF<sub>2</sub>CF<sub>2</sub>H, COEt, cyclopropyl, CF<sub>2</sub>CF<sub>3</sub>, CH=CHCN, allyl, azido, OCF<sub>3</sub>, OCHF<sub>2</sub>, O-i-Pr, SCN, SCHF<sub>2</sub>, SOME, NH-CN, or joined with R<sup>3</sup> and the phenyl carbons to which R<sup>2</sup> and R<sup>3</sup> are attached to form an ethylenedioxy, a dihydrofuryl ring with the oxygen adjacent to a phenyl carbon, or a dihydropyryl ring with the oxygen adjacent to a phenyl carbon;  
 R<sup>3</sup> is H, Et, or joined with R<sup>2</sup> and the phenyl carbons to which R<sup>2</sup> and R<sup>3</sup> are attached to form an ethylenedioxy, a dihydrofuryl ring with the oxygen adjacent to a phenyl carbon, or a dihydropyryl ring with the oxygen adjacent to a phenyl carbon;  
 R<sup>4</sup>, R<sup>5</sup>, and R<sup>6</sup> are independently H, Me, Et, F, Cl, Br, formyl, CF<sub>3</sub>, CHF<sub>2</sub>, CHCl<sub>2</sub>, CH<sub>2</sub>F, CH<sub>2</sub>Cl, CH<sub>2</sub>OH, CN, C<sup>o</sup>CH, 1-propynyl, 2-propynyl, vinyl, OMe, OEt, SMe, or SEt.

[0203] Applicants' invention provides for modulation of gene expression in prokaryotic and eukaryotic host cells. Thus, the present invention also relates to a method for modulating gene expression in a host cell selected from the group consisting of a bacterial cell, a fungal cell, a yeast cell, a plant cell, an animal cell, and a mammalian cell. Preferably, the host cell is a yeast cell, a plant cell, a murine cell, or a human cell.

[0204] Expression in transgenic host cells may be useful for the expression of various polypeptides of interest including but not limited to therapeutic polypeptides, pathway intermediates; for the modulation of pathways already existing in the host for the synthesis of new products heretofore not possible using the host; cell based assays; and the like. Additionally the gene products may be useful for conferring higher growth yields of the host or for enabling alternative growth mode to be utilized.

#### HOST CELLS AND NON-HUMAN ORGANISMS OF THE INVENTION

[0205] As described above, the gene expression modulation system of the present invention may be used to modulate gene expression in a host cell. Expression in transgenic host cells may be useful for the expression of various genes of interest. Thus, Applicants' invention also provides an isolated host cell comprising a gene expression system according to the invention. The present invention also provides an isolated host cell comprising a gene expression cassette according to the invention. Applicants' invention also provides an isolated host cell comprising a polynucleotide or polypeptide according to the invention. The isolated host cell may be either a prokaryotic or a eukaryotic host cell.

[0206] Preferably, the host cell is selected from the group consisting of a bacterial cell, a fungal cell, a yeast cell, a plant cell, an animal cell, and a mammalian cell. Examples of preferred host cells include, but are not limited to, fungal or yeast species such as *Aspergillus*, *Trichoderma*, *Saccharomyces*, *Pichia*, *Candida*, *Hansenula*, or bacterial species such as those in the genera *Synechocystis*, *Synechococcus*, *Salmonella*, *Bacillus*, *Acinetobacter*, *Rhodococcus*, *Streptomyces*, *Escherichia*, *Pseudomonas*, *Methylomonas*, *Methylobacter*, *Alcaligenes*, *Synechocystis*, *Anabaena*, *Thiobacillus*, *Methanobacterium* and *Klebsiella*, plant, animal, and mammalian host cells. More preferably, the host cell is a yeast cell, a plant cell, a mucine cell, or a human cell.

[0207] In a specific embodiment, the host cell is a yeast cell selected from the group consisting of a *Saccharomyces*, a *Pichia*, and a *Candida* host cell.

[0208] In another specific embodiment, the host cell is a plant cell selected from the group consisting of an apple, *Arabidopsis*, bajra, banana, barley, bean, beet, blackgram, chickpea, chili, cucumber, eggplant, favabean, maize, melon,

millet, mungbean, oat, okra, *Panicum*, papaya, peanut, pea, pepper, pigeonpea, pineapple, *Phaseolus*, potato, pumpkin, rice, sorghum, soybean, squash, sugarcane, sugarbeet, sunflower, sweet potato, tea, tomato, tobacco, watermelon, and wheat host cell.

[0209] In another specific embodiment, the host cell is a murine cell.

[0210] In another specific embodiment, the host cell is a human cell.

[0211] Host cell transformation is well known in the art and may be achieved by a variety of methods including but not limited to electroporation, viral infection, plasmid/vector transfection, non-viral vector mediated transfection, *Agrobacterium*-mediated transformation, particle bombardment, and the like. Expression of desired gene products involves culturing the transformed host cells under suitable conditions and inducing expression of the transformed gene. Culture conditions and gene expression protocols in prokaryotic and eukaryotic cells are well known in the art (see General Methods section of Examples). Cells may be harvested and the gene products isolated according to protocols specific for the gene product.

[0212] In addition, a host cell may be chosen which modulates the expression of the inserted polynucleotide, or modifies and processes the polypeptide product in the specific fashion desired. Different host cells have characteristic and specific mechanisms for the translational and post-translational processing and modification (e.g., glycosylation, cleavage [e.g., of signal sequence]) of proteins. Appropriate cell lines or host systems can be chosen to ensure the desired modification and processing of the foreign protein expressed. For example, expression in a bacterial system can be used to produce a non-glycosylated core protein product. However, a polypeptide expressed in bacteria may not be properly folded. Expression in yeast can produce a glycosylated product. Expression in eukaryotic cells can increase the likelihood of "native" glycosylation and folding of a heterologous protein. Moreover, expression in mammalian cells can provide a tool for reconstituting, or constituting, the polypeptide's activity. Furthermore, different vector/host expression systems may affect processing reactions, such as proteolytic cleavages, to a different extent.

[0213] Applicants' invention also relates to a non-human organism comprising an isolated host cell according to the invention. Preferably, the non-human organism is selected from the group consisting of a bacterium, a fungus, a yeast, a plant, an animal, and a mammal. More preferably, the non-human organism is a yeast, a plant, a mouse, a rat, a rabbit, a cat, a dog, a bovine, a goat, a pig, a horse, a sheep, a monkey, or a chimpanzee.

[0214] In a specific embodiment, the non-human organism is a yeast selected from the group consisting of *Saccharomyces*, *Pichia*, and *Candida*.

[0215] In another specific embodiment, the non-human organism is a plant selected from the group consisting of an apple, *Arabidopsis*, bajra, banana, barley, beans, beet, blackgram, chickpea, chili, cucumber, eggplant, fava bean, maize, melon, millet, mungbean, oat, okra, *Panicum*, papaya, peanut, pea, pepper, pigeonpea, pineapple, *Phaseolus*, potato, pumpkin, rice, sorghum, soybean, squash, sugarcane, sugarbeet, sunflower, sweet potato, tea, tomato, tobacco, watermelon, and wheat.

[0216] In another specific embodiment, the non-human organism is a *Mus musculus* mouse.

## MEASURING GENE EXPRESSION/TRANSCRIPTION

[0217] One useful measurement of Applicants' methods of the invention is that of the transcriptional state of the cell including the identities and abundances of RNA, preferably mRNA species. Such measurements are conveniently conducted by measuring cDNA abundances by any of several existing gene expression technologies.

[0218] Nucleic acid array technology is a useful technique for determining differential mRNA expression. Such technology includes, for example, oligonucleotide chips and DNA microarrays. These techniques rely on DNA fragments or oligonucleotides which correspond to different genes or cDNAs which are immobilized on a solid support and hybridized to probes prepared from total mRNA pools extracted from cells, tissues, or whole organisms and converted to cDNA. Oligonucleotide chips are arrays of oligonucleotides synthesized on a substrate using photolithographic techniques. Chips have been produced which can analyze for up to 1700 genes. DNA microarrays are arrays of DNA samples, typically PCR products, that are robotically printed onto a microscope slide. Each gene is analyzed by a full or partial-length target DNA sequence. Microarrays with up to 10,000 genes are now routinely prepared commercially. The primary difference between these two techniques is that oligonucleotide chips typically utilize 25-mer oligonucleotides which allow fractionation of short DNA molecules whereas the larger DNA targets of microarrays, approximately 1000 base pairs, may provide more sensitivity in fractionating complex DNA mixtures.

[0219] Another useful measurement of Applicants' methods of the invention is that of determining the translation state of the cell by measuring the abundances of the constituent protein species present in the cell using processes well known in the art.

[0220] Where identification of genes associated with various physiological functions is desired, an assay may be employed in which changes in such functions as cell growth, apoptosis, senescence, differentiation, adhesion, binding to a specific molecule, binding to another cell, cellular organization, organogenesis, intracellular transport, transport facilitation, energy conversion, metabolism, myogenesis, neurogenesis, and/or hematopoiesis is measured.

[0221] In addition, selectable marker or reporter gene expression may be used to measure gene expression modulation using Applicants' invention.

[0222] Other methods to detect the products of gene expression are well known in the art and include Southern blots (DNA detection), dot or slot blots (DNA, RNA), Northern blots (RNA), and RT-PCR (RNA) analyses. Although less preferred, labeled proteins can be used to detect a particular nucleic acid sequence to which it hybridizes.

[0223] In some cases it is necessary to amplify the amount of a nucleic acid sequence. This may be carried out using one or more of a number of suitable methods including, for example, polymerase chain reaction ("PCR"), ligase chain reaction ("LCR"), strand displacement amplification ("SDA"), transcription-based amplification, and the like. PCR is carried out in accordance with known techniques in which, for example, a nucleic acid sample is treated in the presence of a heat stable DNA polymerase, under hybridizing conditions, with one oligonucleotide primer for each strand of the specific sequence to be detected. An extension product of each primer that is synthesized is complementary to each of the two nucleic acid strands, with the primers sufficiently complementary to each strand of the specific sequence to hybridize therewith. The extension product synthesized from each primer can also serve as a template for further synthesis of extension products using the same primers. Following a sufficient number of rounds of synthesis of extension products, the sample may be analyzed as described above to assess whether the sequence or sequences to be detected are present.

[0224] The present invention may be better understood by reference to the following nonlimiting Examples, which are provided as exemplary of the invention.

## EXAMPLES

### GENERAL METHODS

[0225] Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described by Sambrook, J., Fritsch, E. F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press; Cold Spring Harbor, (1989) (Maniatis) and by T. J. Silhavy, M. L. Bannan, and L. W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1984) and by Ausubel, F. M. et al., *Current Protocols in Molecular Biology*, Greene Publishing Assoc. and Wiley-Interscience (1987).

[0226] Methods for plant tissue culture, transformation, plant molecular biology, and plant, general molecular biology may be found in *Plant Tissue Culture Concepts and Laboratory Exercises* edited by RN Trigiano and DJ Gray, 2<sup>nd</sup> edition, 2000, CRC press, New York; *Agrobacterium Protocols* edited by KMA Gartland and MR Davey, 1995, Humana Press, Totowa, New Jersey; *Methods in Plant Molecular Biology*, P. Maliga et al., 1995, Cold Spring Harbor Lab Press, New York; and *Molecular Cloning*, J. Sambrook et al., 1989, Cold Spring Harbor Lab Press, New York.

[0227] Materials and methods suitable for the maintenance and growth of bacterial cultures are well known in the art. Techniques suitable for use in the following examples may be found as set out in *Manual of Methods for General Bacteriology* (Philipp Gerhardt, R. G. E. Murray, Ralph N. Costilow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg and G. Briggs Phillips, eds), American Society for Microbiology, Washington, DC. (1994) or by Thomas D. Brock in *Biotechnology: A Textbook of Industrial Microbiology*, Second Edition, Sinauer Associates, Inc., Sunderland, MA (1989). All reagents, restriction enzymes and materials used for the growth and maintenance of host cells were obtained from Aldrich Chemicals (Milwaukee, WI), DIFCO Laboratories (Detroit, MI), GIBCO/BRL (Gaithersburg, MD), or Sigma Chemical Company (St. Louis, MO) unless otherwise specified.

[0228] Manipulations of genetic sequences may be accomplished using the suite of programs available from the Genetics Computer Group Inc. (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, WI). Where the GCG program "Pileup" is used the gap creation default value of 12, and the gap extension default value of 4 may be used. Where the GCG "Gap" or "Bestfit" programs is used the default gap creation penalty of 50 and the default gap extension penalty of 3 may be used. In any case where GCG program parameters are not prompted for, in these or any other GCG program, default values may be used.

[0229] The meaning of abbreviations is as follows: "h" means hour(s), "min" means minute(s), "sec" means second(s), "d" means day(s), "μl" means microliter(s), "ml" means milliliter(s), "L" means liter(s), "μM" means micromolar, "mM" means millimolar, "μg" means microgram(s), "mg" means milligram(s), "A" means adenine or adenosine, "T" means thymine or thymidine, "G" means guanine or guanosine, "C" means cytosine or cytosine, "x g" means times gravity, "nt" means nucleotide(s), "aa" means amino acid(s), "bp" means base pair(s), "kb" means kilobase(s), "k" means kilo, "μ" means micro, and "°C" means degrees Celsius.

## EXAMPLE 1

[0230] Applicants' improved EcR-based inducible gene modulation system was developed for use in various applications including gene therapy, expression of proteins of interest in host cells, production of transgenic organisms, and

cell-based assays. This Example describes the construction and evaluation of several gene expression cassettes for use in the EcR-based inducible gene expression system of the invention.

**[0231]** In various cellular backgrounds, including mammalian cells, insect ecdysone receptor (EcR) heterodimerizes with retinoid X receptor (RXR) and, upon binding of ligand, transactivates genes under the control of ecdysone response elements. Applicants constructed several EcR-based gene expression cassettes based on the spruce budworm *Choristoneura fumiferana* EcR ("CfEcR"; full length polynucleotide and amino acid sequences are set forth in SEQ ID NO: 49 and SEQ ID NO: 50, respectively), *C. fumiferana* ultraspiracle ("CfUSP"; full length polynucleotide and amino acid sequences are set forth in SEQ ID NO: 51 and SEQ ID NO: 52, respectively), and mouse *Mus musculus* RXR $\alpha$  (MmRXR $\alpha$ ; full length polynucleotide and amino acid sequences are set forth in SEQ ID NO: 53 and SEQ ID NO: 54, respectively). The prepared receptor constructs comprise a ligand binding domain of EcR and of RXR or of USP; a DNA binding domain of GAL4 or of EcR; and an activation domain of VP16. The reporter constructs include a reporter gene, luciferase or LacZ, operably linked to a synthetic promoter construct that comprises either GAL4 or EcR/USP binding sites (response elements). Various combinations of these receptor and reporter constructs were cotransfected into CHO, NIH3T3, CV1 and 293 cells. Gene induction potential (magnitude of induction) and ligand specificity and sensitivity were examined using four different ligands: two steroidal ligands (ponasterone A and muristerone A) and two non-steroidal ligands (N-(2-ethyl-3-methoxybenzoyl)-N'-(3,5-dimethylbenzoyl)-N'-tert-butylhydrazine and N-(3,4-(1,2-ethylenedioxy)-2-methylbenzoyl)-N-(3,5-dimethylbenzoyl)-N'-tert-butylhydrazine) in a dose-dependent induction of reporter gene expression in the transfected cells. Reporter gene expression activities were assayed at 24hr or 48hr after ligand addition.

**Gene Expression Cassettes:** Ecdysone receptor-based, chemically inducible gene expression cassettes (switches) were constructed as followed, using standard cloning methods available in the art. The following is brief description of preparation and composition of each switch.

**1.1 - GAL4EcR/VP16RXR:** The D, E, and F domains from spruce budworm *Choristoneura fumiferana* EcR ("CfEcRDEF"; SEQ ID NO: 3) were fused to GAL4 DNA binding domain ("DNABD"; SEQ ID NO: 41) and placed under the control of an SV40e promoter (SEQ ID NO: 55). The DEF domains from mouse (*Mus musculus*) RXR ("MmRXRDEF"; SEQ ID NO: 22) were fused to the activation domain from VP16 ("VP16AD"; SEQ ID NO: 45) and placed under the control of an SV40e promoter (SEQ ID NO: 55). Five consensus GAL4 binding sites ("5XGAL4RE"; comprising 5, GAL4RE comprising SEQ ID NO: 47) were fused to a synthetic Et1b minimal promoter (SEQ ID NO: 56) and placed upstream of the luciferase gene (SEQ ID NO: 57).

**1.2 - GAL4EcR/VP16USP:** This construct was prepared in the same way as in switch 1.1 above except MmRXRDEF was replaced with the D, E and F domains from spruce budworm USP ("CfUSPDEF"; SEQ ID NO: 58). The constructs used in this example are similar to those disclosed in U. S. Patent No. 5,880,333 except that *Choristoneura fumiferana* USP rather than *Drosophila melanogaster* USP was utilized.

**1.3 - GAL4RXR/VP16CfEcR:** MmRXRDEF (SEQ ID NO: 22) was fused to a GAL4DNABD (SEQ ID NO: 41) and CfEcRDEF (SEQ ID NO: 1) was fused to a VP16AD (SEQ ID NO: 45).

**1.4 - GAL4RXR/VP16DmEcR:** This construct was prepared in the same way as switch 1.3 except CfEcRDEF was replaced with DmEcRDEF (SEQ ID NO: 6).

**1.5 - GAL4USP/VP16CfEcR:** This construct was prepared in the same way as switch 1.3 except MmRXRDEF was replaced with CfUSPDEF (SEQ ID NO: 58).

**1.6 - GAL4RXRCfEcRVp16:** This construct was prepared so that both the GAL4 DNABD and the VP16AD were placed on the same molecule. GAL4DNABD (SEQ ID NO: 41) and VP16AD (SEQ ID NO: 45) were fused to CfEcRDEF (SEQ ID NO: 3) at N- and C-termini respectively. The fusion was placed under the control of an SV40e promoter (SEQ ID NO: 55).

**1.7 - VP16CfEcR:** This construct was prepared such that CfEcRDEF (SEQ ID NO: 1) was fused to VP16AD (SEQ ID NO: 45) and placed under the control of an SV40e promoter (SEQ ID NO: 55). Six ecdysone response elements ("EcRE"; SEQ ID NO: 59) from the hsp27 gene were placed upstream of the promoter and a luciferase gene (SEQ ID NO: 57). This switch most probably uses endogenous RXR.

**1.8 - DmVgRXR:** This system was purchased from Invitrogen Corp., Carlsbad, California. It comprises a *Drosophila melanogaster* EcR ("DmEcR") with a modified DNABD fused to VP16AD and placed under the control of a CMV promoter (SEQ ID NO: 60). Full length MmRXR (SEQ ID NO: 53) was placed under the control of the RSV promoter (SEQ ID NO: 61). The reporter, pIND(SP1)LacZ, contains five copies of a modified ecdysone response element ("EcRE", E/GRE), three copies of an SP1 enhancer, and a minimal heat shock promoter, all of which were placed upstream of the LacZ reporter gene.

**1.9 - CfVgRXR:** This example was prepared in the same way as switch 1.8 except DmEcR was replaced with a truncated CfEcR comprising a partial A/B domain and the complete CDEF domains [SEQ ID NO: 62 (polynucleotide) and SEQ ID NO: 63 (polypeptide)].

**1.10 - CfVgRXRdel:** This example was prepared in the same way as switch 1.9 except MmRXR (SEQ ID NO:

53) was deleted.

Cell lines: Four cell lines: CHO, Chinese hamster *Cricetulus griseus* ovarian cell line; NIH3T3 (3T3) mouse *Mus musculus* cell line; 293 human *Homo sapiens* kidney cell line, and CV1 African green monkey kidney cell line were used in these experiments. Cells were maintained in their respective media and were subcultured when they reached 60% confluency. Standard methods for culture and maintenance of the cells were followed.

Transfections: Several commercially available lipofactors as well as electroporation methods were evaluated and the best conditions for transfection of each cell line were developed. CHO, NIH3T3, 293 and CV1 cells were grown to 60% confluency. DNAs corresponding to the various switch constructs outlined in Examples 1.1 through 1.10 were transfected into CHO cells, NIH3T3 cells, 293 cells, or CV1 cells as follows.

CHO cells: Cells were harvested when they reach 60-80% confluency and plated in 6- or 12- or 24- well plates at 250,000, 100,000, or 50,000 cells in 2.5, 1.0, or 0.5 ml of growth medium containing 10% Fetal bovine serum respectively. The next day, the cells were rinsed with growth medium and transfected for four hours. LipofectAMINE™ 2000 (Life Technologies Inc.) was found to be the best transfection reagent for these cells. For 12- well plates, 4  $\mu$ l of LipofectAMINE™ 2000 was mixed with 100  $\mu$ l of growth medium 1.0  $\mu$ g of reporter construct and 0.25  $\mu$ g of receptor construct(s) were added to the transfection mix. A second reporter construct was added [pTKRL (Promega), 0.1  $\mu$ g/transfection mix] and comprised a *Renilla luciferase* gene (SEQ ID NO: 64) operably linked and placed under the control of a thymidine kinase (TK) constitutive promoter and was used for normalization. The contents of the transfection mix were mixed in a vortex mixer and let stand at room temperature for 30 min. At the end of incubation, the transfection mix was added to the cells maintained in 400  $\mu$ l growth medium. The cells were maintained at 37 °C and 5% CO<sub>2</sub> for four hours. At the end of incubation, 500  $\mu$ l of growth medium containing 20% FBS and either DMSO (control) or a DMSO solution of appropriate ligands were added and the cells were maintained at 37 °C and 5% CO<sub>2</sub> for 24-48 hr. The cells were harvested and reporter activity was assayed. The same procedure was followed for 6 and 24 well plates as well except all the reagents were doubled for 6 well plates and reduced to half for 24-well plates.

NIH3T3 Cells: Superfect™ (Qiagen Inc.) was found to be the best transfection reagent for 3T3 cells. The same procedures described for CHO cells were followed for 3T3 cells as well with two modifications. The cells were plated when they reached 50% confluency. 125,000 or 50,000 or 25,000 cells were plated per well of 6- or 12- or 24-well plates respectively. The GA14EcR/VP16RXR and reporter vector DNAs were transfected into NIH3T3 cells, the transfected cells were grown in medium containing PonA, MuRA, N-(2-ethyl-3-methoxybenzoyl)-N'-(3,5-dimethylbenzoyl)-N'-t-butylhydrazine, or N-(3-(4-(1,2-ethylenedioxy)-2-methylbenzoyl)-N-(3,5-dimethylbenzoyl)-N'-tert-butylhydrazine for 48 hr. The ligand treatments were performed as described in the CHO cell section above.

293 Cells: LipofectAMINE™ 2000 (Life Technologies) was found to be the best lipofactor for 293 cells. The same procedures described for CHO were followed for 293 cells except that the cells were plated in biocoated plates to avoid clumping. The ligand treatments were performed as described in the CHO cell section above.

CV1 Cells: LipofectAMINE™ plus (Life Technologies) was found to be the best lipofactor for CV1 cells. The same procedures described for NIH3T3 cells were followed for CV1 cells

Ligands: Ponasterone A and Muriesterone A were purchased from Sigma Chemical Company. The two non-steroids N-(2-ethyl-3-methoxybenzoyl)-N'-(3,5-dimethylbenzoyl)-N'-t-butylhydrazine, or N-(3,4-(1,2-ethylenedioxy)-2-methylbenzoyl)-N'-(3,5-dimethylbenzoyl)-N'-tert-butylhydrazine are synthetic stable ecdysteroids synthesized at Rohm and Haas Company. All ligands were dissolved in DMSO and the final concentration of DMSO was maintained at 0.1 % in both controls and treatments.

Reporter Assays: Cells were harvested 24-48 hr after adding ligands. 125, 250, or 500  $\mu$ l of passive lysis buffer (part of Dual-luciferase™ reporter assay system from Promega Corporation) were added to each well of 24- or 12- or 24-well plates respectively. The plates were placed on a rotary shaker for 15 min. Twenty  $\mu$ l of lysate was assayed. Luciferase activity was measured using Dual-luciferase™ reporter assay system from Promega Corporation following the manufacturer's instructions.  $\beta$ -Galactosidase was measured using Galacto-Star™ assay kit from TROPIX following the manufacturer's instructions. All luciferase and  $\beta$ -galactosidase activities were normalized using *Renilla luciferase* as a standard. Fold activities were calculated by dividing normalized relative light units ("RLU") in ligand treated cells with normalized RLU in DMSO treated cells (untreated control).

[0232] The results of these experiments are provided in the following tables.

Table 1

Transactivation of reporter genes through various switches in CHO cells		
Composition of Switch	Mean Fold Activation with 50 $\mu$ M N-(2-ethyl-3-methoxybenzoyl)-N'-(3,5-dimethylbenzoyl)-N'-t-butylhydrazine	
1.1 GAL4EcR + VP16RXR pGAL4RELuc	267	
1.2 GAL4EcR + VP16USP pGAL4RELuc	2	
1.3 GAL4RXR + VP16CfEcR pGAL4RELuc	85	
1.4 GAL4RXR+VP16DmEcR pGAL4RELuc	312	
1.5 GAL4USP + VP16CfEcR pGAL4RELuc	2	
1.6 GAL4CfEcRVP16 pGAL4RELuc	9	
1.7 VP16CfEcR pEcRELuc	36	
1.8 DmVgRXR+MmRXR pIND(SP1)LacZ	14	
1.9 CfVgRXR + MmRXR pIND(SP1)LacZ	27	
1.10 CfVgRXR pIND(SP1)LacZ	29	

Table 2

Transactivation of reporter genes through various switches in 3T3 cells		
Composition of Switch	Mean Fold Activation Through N-(2-ethyl-3-methoxybenzoyl)-N'-(3,5-dimethylbenzoyl)-N'-t-butylhydrazine	
1.1 GAL4EcR + VP16RXR pGAL4RELuc	1118	
1.2 GAL4EcR + VP16USP pGAL4RELuc	2	
1.3 GAL4RXR + VP16CfEcR pGAL4RELuc	47	
1.4 GAL4RXR + VP16DmEcR pGAL4RELuc	269	
1.5 GAL4USP + VP16CfEcR pGAL4RELuc	3	
1.6 GAL4CfEcRVP16 pGAL4RELuc	7	
1.7 VP16CfEcR pEcRELuc	1	
1.8 DmVgRXR+MmRXR pIND(SP1)LacZ	21	
1.9 CfVgRXR + MmR7Gt pIND(SP1)LacZ	19	
1.10 CfVgRXR	2	

(continued)

Transactivation of reporter genes through various switches in 3T3 cells	
Composition of Switch	Mean Fold Activation Through N-(2-ethyl-3-methoxybenzoyl)-N'-(3,5-dimethylbenzoyl)-N'-t-butylhydrazine
pIND(SP1)LacZ	

Table 3

Transactivation of reporter genes through various switches in 293 cells	
Composition of Switch	Mean Fold Activation Through N-(2-ethyl-3-methoxybenzoyl)-N'-(3,5-dimethylbenzoyl)-N'-t-butylhydrazine
1.1 GAL4EcR+VP16RXR pGAL4RELuc	125
1.2 GAL4EcR + VP16USP pGAL4RELuc	2
1.3 GAL4RXR+VP16CFecR pGAL4RELuc	17
1.4 GAL4RXR + VP16DmEcR pGAL4RELuc	3
1.5 GAL4USP + VP16CFecR pGAL4RELuc	2
1.6 GAL4CFecRV16 pGAL4RELuc	3
1.7 VP16CFecR pEcRELuc	2
1.8 DmVgRXR+ MmRXR pIND(SP1)LacZ	21
1.9 CFvRXR + MmRXR pIND(SP1)LacZ	12
1.10 CFvRXR pIND(SP1)LacZ	3

Table 4

Transactivation of reporter genes through various switches in CV1 cells	
Composition of Switch	Mean Fold Activation Through N-(2-ethyl-3-methoxybenzoyl)-N'-(3,5-dimethylbenzoyl)-N'-t-butylhydrazine
1.1 GAL4EcR + VP16RXR pGAL4RELuc	279
1.2 GAL4EcR+VP16USP pGAL4RELuc	2
1.3 GAL4RXR + VP16CFecR pGAL4RELuc	25
1.4 GAL4RXR + VP16DmEcR pGAL4RELuc	80
1.5 GAL4USP + VP16CFecR pGAL4RELuc	3
1.6 GAL4CFecRV16 pGAL4RELuc	6



(continued)

Transactivation of reporter genes through various switches in CV1 cells	
Composition of Switch	Mean Fold Activation Through N-(2-ethyl-3-methoxybenzoyl)-N'-(3,5-dimethylbenzoyl)-N'-tert-butylhydrazine
1.7 VP16CfEcR pEcRELuc	1
1.8 DmVgRXR+MmRXR	12
pIND(SP1)LacZ	
1.9 CfVgRXR + MmRXR	7
pIND(SP1)LacZ	
1.10 Cf VgRXR	1
pIND(SP1)LacZ	

Table 5

Transactivation of reporter gene GAL4CfEcRDEF/VP16MmRXRDEF (switch 1.1) through steroids and non-steroids in 3T3 cells.	
Ligand	Mean Fold Induction at 1.0 $\mu$ M Concentration
1. Ponasterone A	1.0
2. Muristerone A	1.0
3. N-(2-ethyl-3-methoxybenzoyl)-N'-(3,5-dimethylbenzoyl)-N'-tert-butylhydrazine	116
4. N'-(3,4-(1,2-ethylenedioxy)-2-methylbenzoyl)-N'-(3,5-diethylbenzoyl)-N'-tert-butylhydrazine	601

Table 6

Transactivation of reporter gene GAL4MnRXRDEF/VP16CfEcRCDEF (switch 1.3) through steroids and non-steroids in 3T3 cells.	
Ligand	Mean Fold Induction at 1.0 $\mu$ M Concentration
1. Ponasterone A	1.0
2. Muristerone A	1.0
3. N-(2-ethyl-3-methoxybenzoyl)-N'-(3,5-dimethylbenzoyl)-N'-tert-butylhydrazine	71
4. N'-(3,4-(1,2-ethylenedioxy)-2-methylbenzoyl)-N'-(3,5-dimethylbenzoyl)-N'-tert-butylhydrazine	54

[0233] Applicants' results demonstrate that the non-steroidal ecdysone agonists, N-(2-ethyl-3-methoxybenzoyl)-N'-(3,5-dimethylbenzoyl)-N'-tert-butylhydrazine and N'-(3,4-(1,2-ethylenedioxy)-2-methylbenzoyl)-N'-(3,5-dimethylbenzoyl)-N'-tert-butylhydrazine, were more potent activators of CfEcR as compared to *Drosophila melanogaster* EcR (DmEcR). (see Tables 1-4). Also, in the mammalian cell lines tested, MmRXR performed better than CfUSP as a heterodimeric partner for CfEcR. (see Tables 1-4). Additionally, Applicants' inducible gene expression modulation system performed better when exogenous MmRXR was used than when the system relied only on endogenous RXR levels (see Tables 1-4).

[0234] Applicants' results also show that in a CfEcR-based inducible gene expression system, the non-steroidal ecdysone agonists induced reporter gene expression at a lower concentration (i.e., increased ligand sensitivity) as compared to the steroid ligands, ponasterone A and muristerone A (see Tables 5 and 6).

[0235] Out of 10 EcR based gene switches tested, the GAL4EcR/VP16RXR switch (Switch 1.1) performed better than any other switch in all four cell lines examined and was more sensitive to non-steroids than steroids. The results also demonstrate that placing the activation domain (AD) and DNA binding domain (DNABD) on each of the two partners reduced background when compared to placing both AD and DNABD together on one of the two partners. Therefore,

a switch format where the AD and DNABD are separated between two partners, works well for EcR-based gene switch applications.

[0236] In addition, the MmRXR/EcR-based switches performed better than CfUSP/EcR-based switches, which have a higher background activity than the MmRXR/EcR switches in the absence of ligand.

[0237] Finally, the GAL4EcR/VP16RXR switch (Switch 1.1) was more sensitive to non-steroid ligands than to the steroid ligands (see Tables 5 and 6). In particular, steroid ligands initiated transactivation at concentrations of 50  $\mu$ M, whereas the non-steroid ligands initiated transactivation at less than 1  $\mu$ M (submicromolar) concentration.

## EXAMPLE 2

[0238] This Example describes Applicants' further analysis of truncated EcR and RXR polypeptides in the improved EcR-based inducible gene expression system of the invention. To identify the best combination and length of two receptors that give a switch with a) maximum induction in the presence of ligand; b) minimum background in the absence of ligand; c) highly sensitive to ligand concentration; and d) minimum cross-talk among ligands and receptors, Applicants made and analyzed several truncation mutations of the CfEcR and MmRXR receptor polypeptides in NIH3T3 cells.

[0239] Briefly, polynucleotides encoding EcR or RXR receptors were truncated at the junctions of A/B, C, D, E and F domains and fused to either a GAL4 DNA binding domain encoding polynucleotide (SEQ ID NO: 41) for CfEcR, or a VP16 activation domain encoding polynucleotide (SEQ ID NO: 45) for MmRXR as described in Example 1. The resulting receptor truncation/fusion polypeptides were assayed in NIH3T3 cells. Plasmid pFRLUC (Stratagene) encoding a luciferase polypeptide was used as a reporter gene construct and pTKRL (Promega) encoding a *Renilla* luciferase polypeptide under the control of the constitutive TK promoter was used to normalize the transfections as described above. The analysis was performed in triplicates and mean luciferase counts were determined as described above.

### Gene Expression Cassettes Encoding Truncated Ecdysone Receptor Polypeptides

[0240] Gene expression cassettes comprising polynucleotides encoding either full length or truncated CfEcR polypeptides fused to a GAL4 DNA binding domain (SEQ ID NO: 41): GAL4CfEcRABGDEF (full length CfEcR/BCDEF; SEQ ID NO: 49), GAL4CfEcRCDEF (CfEcRCDEF; SEQ ID NO: 1), GAL4CfEcR1/2CDEF (CfEcR1/2CDEF; SEQ ID NO: 2), GAL4CfEcRDEF (CfEcRDEF; SEQ ID NO: 3), GAL4CfEcREF (CfEcREF; SEQ ID NO: 4), and GAL4CfEcRDE (CfEcRDE; SEQ ID NO: 5) were transfected into NIH3T3 cells along with VP16MmRXRDEF (constructed as in Example 1.1; Figure 11) or VP16MmRXREF (constructed as in Example 1.1 except that MmRXRDEF was replaced with MmRXREF (SEQ ID NO: 23); Figure 12), and pFRLUC and pTKRL plasmid DNAs. The transfected cells were grown in the presence of 0.1, 5 or 25  $\mu$ M of N-(2-ethyl-3-methoxybenzoyl)-N'-(3,5-dimethylbenzoyl)-N'-tert-butylhydrazine or PonA for 48 hr. The cells were harvested, lysed and luciferase reporter activity was measured in the cell lysates. Total fire luciferase relative light units are presented. The number on the top of each bar is the maximum fold induction for that treatment.

[0241] Applicants' results show that the EF domain of MmRXR is sufficient and performs better than DEF domains of this receptor (see Figures 11 and 12). Applicants have also shown that, in general, EcR/RXR receptor combinations are insensitive to PonA (see Figures 11 and 12). As shown in the Figures 11 and 12, the GAL4CfEcRCDEF hybrid polypeptide (SEQ ID NO: 7) performed better than any other CfEcR hybrid polypeptide.

### Gene Expression Cassettes Encoding Truncated Retinoid X Receptor Polypeptides

[0242] Gene expression cassettes comprising polynucleotides encoding either full length or truncated MmRXR polypeptides fused to a VP16 transactivation domain (SEQ ID NO: 45): VP16MmRXRA/BCDEF (full length MmRXRA/BCDEF; SEQ ID NO: 53), VP16MmRXRCDEF (MmRXRCDEF; SEQ ID NO: 21), VP16MmRXRDEF (MmRXRDEF; SEQ ID NO: 22), VP16MmRXREF (MmRXREF; SEQ ID NO: 23), VP16MmRXRBam-EF ("MmRXRBam-EF" or "MmRXR-truncatedEF"; SEQ ID NO: 24), and VP16MmRXRAF2de1 ("MmRXRAF2de1" or "MmRXR-E"; SEQ ID NO: 25) constructs were transfected into NIH3T3 cells along with GAL4CfEcRCDEF (constructed as in Example 1.1; Figure 13) or GAL4CfEcRDEF [constructed as in Example 1.1 except CfEcRCDEF was replaced with CfEcRDEF (SEQ ID NO: 3); Figure 14], pFRLUC and pTKRL plasmid DNAs as described above. The transfected cells were grown in the presence of 0, 1, 5 and 25  $\mu$ M of N-(2-ethyl-3-methoxybenzoyl)-N'-(3,5-dimethylbenzoyl)-N'-tert-butylhydrazine or PonA for 48 hr. The cells were harvested and lysed and reporter activity was measured in the cell lysate. Total fire luciferase relative light units are presented. The number on top of each bar is the maximum fold induction in that treatment.

[0243] Of all the truncations of MmRXR tested, Applicants' results show that the MmRXREF receptor was the best partner for CfEcR (Figures 13 and 14). CfEcRCDEF showed better induction than CfEcRDEF using MmRXREF. Deleting AF2 (abbreviated "EF-AF2de1") or helices 1-3 of the E domain (abbreviated "EF-Bamdef") resulted in an RXR receptor that reduced gene induction and ligand sensitivity when partnered with either CfEcRCDEF (Figure 13) or CfEcRDEF (Figure 14) in NIH3T3 cells. In general, the CfEcR/RXR-based switch was much more sensitive to the non-steroid N-(2-

ethyl-3-methoxybenzoyl)-N<sup>-</sup>-(3,5-dimethylbenzoyl)-N<sup>-</sup>-tert-butylhydrazine than to the steroid PonA.

### EXAMPLE 3

[0244] This Example describes Applicants' further analysis of gene expression cassettes encoding truncated EcR or RXR receptor polypeptides that affect either ligand binding activity or ligand sensitivity, or both. Briefly, six different combinations of chimeric receptor pairs, constructed as described in Examples 1 and 2, were further analyzed in a single experiment in NIH3T3 cells. These six receptor pair combinations and their corresponding sample numbers are depicted in Table 7.

Table 7

CfEcR + MmRXR Truncation Receptor Combinations in NIH3T3 Cells		
Figure 15 X-Axis Sample No.	EcR Polypeptide Construct	RXR Polypeptide Construct
Samples 1 and 2	GAL4CfEcRCDEF	VP16RXRABCDEF (Full length)
Samples 3 and 4	GAL4CfEcRCDEF	VP16RXRDEF
Samples 5 and 6	GAL4CfEcRCDEP	VP16RXREF
Samples 7 and 8	GAL4CfEcRDEF	VP16RXRA/BCDEF (Full length)
Samples 9 and 10	GAL4CfEcRDEF	VP16RXRDEF
Samples 11 and 12	GAL4CfEcRDEF	VP16RXREF

[0245] The above receptor construct pairs, along with the reporter plasmid pRLuc were transfected into NIH3T3 cells as described above. The six CfEcR truncation receptor combinations were duplicated into two groups and treated with either steroid (odd numbers on x-axis of Figure 15) or non-steroid (even numbers on x-axis of Figure 15). In particular, the cells were grown in media containing 0, 1, 5 or 25 uM PonA (steroid) or N-(2-ethyl-3-methoxybenzoyl)-N<sup>-</sup>-(3,5-dimethylbenzoyl)-N<sup>-</sup>-tert-butylhydrazine (non-steroid) ligand. The reporter gene activity was measured and total RLU are shown. The number on top of each bar is the maximum fold induction for that treatment and is the mean of three replicates.

[0246] As shown in Figure 15, the CfEcRCDEF/MmRXREF receptor combinations were the best switch pairs both in terms of total RLU and fold induction (compare columns 1-6 to columns 7-12). This confirms Applicants' earlier findings as described in Example 2 (Figures 11-14). The same gene expression cassettes encoding the truncated EcR and RXR polypeptides were also assayed in a human lung carcinoma cell line A549 (ATCC) and similar results were observed (data not shown).

### SEQUENCE LISTING

#### [0247]

```

<110> Rohm and Haas Company
Palli, Subba Reddy
Kapitskaya, Marianna Zinovjevna
Cress, Dean Ervin
<120> Novel Ecdysone Receptor-Based Inducible Gene Expression System
<130> RH0020
<150> 60/191,355
<151> 2000-03-22
<150> 60/269,799
<151> 2001-02-20
<160> 64
<170> PatentIn version 3.0 ,
<210> 1
<211> 1288
<212> DNA
<213> Artificial Sequence
<220>
<221> misc feature

```

# EP 1 266 015 B1

<223> Novel Sequence

<400> 1

5	aagggccctg cgcccgctca gcaagaggaa ctgtgtctgg tatgcgggga cagagcctcc	60
	ggataccact acaatgctgct cacgtgtgaa ggggtgtaaag ggttcctcag acggagtgtt	120
	accaaaaatg cgggtttatat ttgtaaatcc ggtcacgctt gcgaaatgga catgtacatg	180
10	cgacggaat gccaggagtg ccgctctgaag aagtgccttag ctgtaggcat gaggcctgag	240
	tgcgtagtac ccgagactca gtgcgccatg aagcggaaag agaagaagc acgaaggag	300
	aaggacaac tgctgtctcag cagcagcagc gtggacgacc acatgcgcgc cattatgcag	360
	tgtgaacctc cactctctga agcagcaagg attcacgaag tgggtcccaag gttctctctc	420
15	gacaagctgt tggagacaaa ccggcagaaa aacatccccc agttgacagc caaccagcag	480
	ttccttatcg ccaggctcat ctgggtaccg gacgggtacg agcagccttc tgatgaagat	540
	ttgaagagga ttacgcagac gtggcagcaa gcggacgatg aaaacgaaga gtctgacact	600
20	cccttcgcc agatcacaga gatgactatc ctacacgggtc aacttatcgt ggagtctcgc	660
	aagggattgc cagggttcgc caagatctcg cagcctgac aaattacgct gcttaaggct	720
	tgtcaagtgc aggtaatgat gctccgagtc gcgcgacgat acgatgcgcg ctacagacagt	780
25	gttctgttcg cgaacaacca agcgtacact cgcgacaact accgcaaggc tggcatggcc	840
	tacgtcatcg aggatctact gcaactctgc cgggtcatgt actctatggc gttggacaac	900
	atccattacg cgtgtctcac ggctgtctgc atctttctg accggccagg gttggagcag	960
30	ccgcaactg tgaagaaat ccagcggtag tacctgaata cgctccgcat ctatactctg	1020
	aaccagctga ggggtcgcg gcgttcgtcc gtcatatacg gcaagatcct ctcaatctc	1080
35	tctgagctac gcacgctcg catgcaaac tccaacatgt gcattctcct caagctcaag	1140
	aacagaaagc tgccgccttt cctcgaggag atctgggatg tggcgacat gtcgcacacc	1200
	caaccgcgc ctactctcga gtcccccacg aatctctagc cctgcgcgc acgcatcgcc	1260
40	gatgccgcgt ccggccgcgc tgctctga	1288

<210> 2

<211> 1110

<212> DNA

<213> Artificial Sequence

<220>

<221> misc\_feature

<223> Novel Sequence

<400> 2

EP 1 266 015 B1

	gcgggtttata tttgtaaatt cggtcacgct tgcgaaatgg acatgtacat gcgacgaaaa	60
	tgccaggagat gccgcctgaa gaagtgcctta gctgtaggca tgaggcctga gtgcgtagta	120
5	cccagagactc agtgcgccat gaagcggaaa gagaagaaa cacagaagga gaaggacaaa	180
	ctgcctgtca gcacgacgac ggtggacgac cacatgccgc ccattatgca gtgtgaacct	240
	ccacctcctg aagcagcaag gattcacgaa gtgggtccaa ggtttctctc cgacaagctg	300
10	ttggagacaa accgcgagaa aaacatcccc cagttgacag ccaaccagca gtctcttctc	360
	gccaggtctc tctggtacca ggacgggtac gacgagcctt ctgatgaaga ttgtaagagg	420
	attacgcaga cgtggcagca agcggacgat gaaaacgaag agtctgacac tcctctccgc	480
15	cagatcacag agatgactat cctcacggtc caacttatcg tggagttcgc gaagggattg	540
	ccagggttgc ccaagatctc gcagcctgat caaattacgc tgcttaaggc ttgctcaagt	600
	gaggtaatga tgctccgagt cgcgcgacga tacgatgcgg cctcagacag tgttctgttc	660
20	gcgaacaacc aagcgtacac tcgcgacaac taccgcaagg ctggcatggc ctacgtcctc	720
	gaggatctac tgcaettctg ccggtgcatg tactctatgg cgttggacaa catccattac	780
	gcgctgctca cggctgtcgt catcttttct gaccggccag ggttggagca gccgcaactg	840
25	gtggaagaaa tccagcggta ctacctgaat acgctccgca tctatatcct gaaccagctg	900
	agcgggtcgg cgcgttctgc cgtcatatac ggcaagatcc tctcaatcct ctctgagcta	960
	cgcacgcctg gcatgcaaaa ctccaacatg tgcatctccc tcaagctcaa gaacagaaag	1020
30	ctgccgcctt tcctcgagga gatctgggat gtggcggaca tgctgcacac ccaaccgcgc	1080
	cctatctcgc agtccccac gaatctctag	1110

<210> 3

<211> 1054

<212> DNA

<213> Artificial Sequence

<220>

<221> misc\_feature

<223> Novel Sequence

<400> 3

EP 1 266 015 B1

5 cctgagtgcg tagtaccga gactcagtc gccatgaagc gaaagagaa gaaagcacag 60  
 aaggagaagg acaaactgcc tgtcagcacg acgacgggtg acgaccacat gccgccatt 120  
 atgcagtggt aacctccacc tccatgaagca gcaaggattc acgaagtgtt cccaaggttt 180  
 ctctccgaca agctgttga gacaaaccgg cagaaaaaca tccccagtt gacagccaac 240  
 cagcagttcc ttatcgccag gctcatctgg taccaggacg ggtacgagca gccttctgat 300  
 10 gaagatttga agaggattac gcagacgtgg cagcaagcgg acgatgaaaa cgaagagtct 360  
 gacactccct tccgccagat cacagagatg actatcctca cggccaact tatcgtggag 420  
 ttccgcaagg gattgccagg gttcgcaag atctcgacg ctgatcaaat tacgctgctt 480  
 15 aaggcttgct caagtgaagt aatgatgctc cgagtcgcgc gacgatacga tgcggcctca 540  
 gacagtgctt tgttcgcgaa caaccaagcg tacactcgcg acaactaccg caaggctggc 600  
 atggcctacg tcatcgagga tctactgcac ttctgccggt gcattgtact tatggcggtt 660  
 20 gacaacatcc attacgcgct gctcacggct gtcgtcatct tttctgacgg gccagggtt 720  
 gagcagccgc aactggtgga agaaatccag cggactactacc tgaatacgcg ccgcatctat 780  
 atcctgaacc agctgagcgg gtcggcgctg tcgtccgtca tatacggcaa gatcctctca 840  
 25 atcctctctg agctacgcac gctcggcgatg caaaactcca acatgtgcat ctccctcaag 900  
 ctcaagaaca gaaagctgcc gcctttcctc gaggagatct gggatgtggc ggacatgtcg 960  
 cacacccaac cgccgcctat cctcgagtcc cccacgaatc tctagcccct gcgcgcacgc 1020  
 atcgccgatg ccgcgtccgg ccgcgctgct ctga 1054

35 <210> 4  
 <211> 735  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <221> misc feature  
 <223> Novel Sequence  
 40 <400> 4

45 taccaggacg ggtacgagca gccttctgat gaagatttga agaggattac gcagacgtgg 60  
 cagcaagcgg acgatgaaaa cgaagagtct gacactccct tccgccagat cacagagatg 120  
 actatcctca cggccaact tatcgtggag ttccgcaagg gattgccagg gttcgccaag 180  
 atctcgacg ctgatcaaat tacgctgctt aaggcttgct caagtgaagt aatgatgctc 240  
 cgagtcgcgc gacgatacga tgcggcctca gacagtgctt tgttcgcgaa caaccaagcg 300

5      tacactcgcg acaactaccg caaggctggc atggcctacg tcatacgagga tctactgcac      360  
 ttctcgccgt gcatgtactc tatggcggtg gacaacatcc attacgcgct gctcaecggct      420  
 gtcgctatct tttctgaccg gccagggttg gagcagccgc aactggtgga agaaaaccag      480  
 cggctactacc tgaatacgcct ccgcactctat atcctgaacc agctgagcgg gtcggcgctg      540  
 tcgtccgtca tatacggcaa gatccctccta atcctctctg agctacgcac gctcggcatg      600  
 10      caaaactcca acatgtgcat ctccctcaag ctcaagaaca gaaagctgcc gcctttcttc      660  
 gaggagatct gggatgtggc ggacatgtcg cacaccaaac cgccgcctat cctcagatcc      720  
 cccacgaatc tctag      735

15      <210> 5  
 <211> 960  
 <212> DNA  
 <213> Artificial Sequence  
 20      <220>  
 <221> misc\_feature  
 <223> Novel Sequence  
 <400> 5

25      cctgagtgcg tagtaccgga gactcagtcg gccatgaagc ggaagagaa gaaagcacag      60  
 aaggagaagg acaaaactgcc tgtcagcacg acgacgggtg acgaccacat gccgccatt      120  
 atgcagtgtg aacctccacc tcctgaagca gcaaggattc acgaagtgtt cccaagggtt      180  
 30      ctctccgaca agctgttgga gacaaaaccg cagaaaaaca tccccagt gacagccaac      240  
 cagcagttec ttatcgccag gctcatctgg taccaggacg ggtacgagca gccttctgat      300  
 gaagatttga agaggattac gcagacgttg cagcaagcgg acgatgaaaa cgaagagtct      360  
 35      gacactccct tcgccagat cacagagatg actatcctca cggtccaact tatcgtggag      420  
 ttctcggaagg gattgccagg gttcgccaag atctcgcagc ctgatcaat tacgctgett      480  
 aaggcttgct caagtggagt aatgatgtct cagtcgcgcg gacgatacga tgcggcctca      540  
 40      gacagtgttc tgttcgcgaa caaccaagcg tacactcgcg acaactaccg caaggctggc      600  
 atggcctacg tcatacgagga tctactgcac ttctgccggt gcatgtactc tatggcggtg      660  
 gacaacatcc attacgcgct gctcacggct gtcgctatct tttctgaccg gccagggttg      720  
 45      gagcagccgc aactggtgga agaaaaccag cggctactacc tgaatacgtc ccgcatctat      780  
 atcctgaacc agctgagcgg gtcggcgctg tcgtccgtca tatacggcaa gatccctccta      840  
 atcctctctg agctacgcac gctcggcatg caaaactcca acatgtgcat ctccctcaag      900  
 50      ctcaagaaca gaaagctgcc gcctttcttc gaggagatct gggatgtggc ggacatgtcg      960

<210> 6  
 <211> 1878  
 <212> DNA  
 55      <213> Artificial Sequence  
 <220>  
 <221> misc feature  
 <223> Novel Sequence

&lt;400&gt; 6

	ggacctgcgc cacgggtgca agaggagctg tgcctgggtt gcggcgacag ggcctccggc	60
5	taccactaca acgccctcac ctgtgagggc tgcaagggtt tctttcgacg cagcggtacg	120
	aagagcgccg tctactgctg caagtccggg cgcgcctgcy aaatggacat gtacatgagg	180
	cgaaagtgtc aggagtgccg cctgaaaaag tgcctggccg tgggtatgcy gccggaatgc	240
10	gtcgtcccg agaaccaatg tgcgatgaag cggcgcaaaa agaaggccca gaaggagaag	300
	gacaaaaatga ccaactcgcc gagctctcag catggcggca atggcagctt ggcctctggt	360
	ggcggccaag actttgttaa gaaggagatt cttgacctta tgacatgcga gccgcccag	420
15	catgccacta ttccgttact acctgatgaa atattggcca agtgtcaagc gcgcaatata	480
	ccttccttaa cgtacaatca gttggccgtt atatacaagt taattggta ccaggatggc	540
	tatgagcagc catctgaaga ggatctcagg cgtataatga gtcaaccgga tgaagaacgag	600
20	agccaaacgg acgtcagctt tcggcatata accgagataa ccatactcac ggtccagttg	660
	attgttgagt ttgctaaagg tctaccagcy ttacaaaaga taccagga ggaccagatc	720
	acgttactaa aggcctgctc gtcggaggtg atgatgctcy gtatggcacg acgctatgac	780
25	cacagctcgg actcaatatt ctccgcaat aatagatcat ataccggga ttcttaca	840
	atggccggaa tggcgtataa cattgaagac ctgctgcatt tctgcccca aatgtctcgc	900
	atgaaggtgy acaacgtcga atacgcgctt ctactgcga ttgtgatctt ctccggaccg	960
	ccgggctcgg agaaggccca actagtcgaa cgcgatccga gctactacat cgacacgcta	1020
30	cgcatttata tactcaaccg ccactgcggc gactcaatga gcctcgtctt ctacgcaaag	1080
	ctgctctcga tctcaccga gctgcgtacg ctgggcaacc agaaccgga gatgtgttc	1140
	tcactaaagc tcaaaaacgg caaactgcc aagttcctcg aggagatctg ggacgttcat	1200
35	gccatccgc catcggtcca gtcgcacett cagattacc aggaggagaa cgagcgtctc	1260
	gagcgggtg agcgtatgcy ggcacggtt gggggcgcca ttaccgcgg cattgattgc	1320
	gactctgcct ccaactcggc ggcggcagcc gcggcccagc atcagcctca gcctcagccc	1380
40	cagcccaccc cctcctccct gaccagaaac gattcccagc accagacaca gccgcagcta	1440
	caacctcagc taccacctca gctgcaaggt caactgcaac ccagctcca accacagctt	1500
	cagacgcaac tccagccaca gattcaacca cagccacagc tcctcccgct ctccgctccc	1560
45	gtcccgctc ccgtaaccgc acctgggttc ttgtccgcy tcagtacgag cagcaataac	1620
	atggggcgaa gtgcggccat aggaacctac acgccggcaa ccaccagcag tattcacgct	1680
	gccgttaccg ctagctccac cacatcagcy gtaccgatgg gcaacggagt tggagtcggt	1740
50		
	gttggggtg gcggcaacgt cagcatgtat gcgaacgcc agacggcgat ggccttgatg	1800
	gggtgagccc tgcattcgca ccaagagcag cttatcgggg gagtggcggt taagtggag	1860
55	cactcgacga ctgcatag	1878

&lt;210&gt; 7



<211> 1752  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <221> misc feature  
 <223> Novel Sequence  
 <400> 7

10	g c g c t c t a c t g c t g c a a g t t c g g g c g c g c t g c g a a a t g g a c a t g t a c a t g a g g c g a a a g	60
	t g t c a g g a g t g c c g c c t g a a a a g t g c c t g g c c g t g g g t a t g c g g c c g g a a t g c g t c g t c	120
	c c g g a g a a c c a a t g t g c g a t g a a g c g g c g c g a a a a g a g g c c c a g a a g g a g a a g a c a a a	180
15	a t g a c c a c t t c g c c g a g c t c t c a g c a t g g c g g c a a t g g c a g c t t g g c c t c t g t g g c g g c	240
	c a a g a c t t t g t t a a g a a g g a g a t t c t t g a c c t t a t g a c a t g c g a g c c g c c c a g c a t g c c	300
	a c t a t t c c g c t a c t a c c t g a t g a a a t a t t g g c a a g t g t c a a g c g c g c a a t a t a c c t t c c	360
20	t t a a c g t a c a t a c a g t t g g c c g t t a t a t a c a a g t t a a t t t g g t a c c a g g a t g g c t a t g a g	420
	c a g c c a t c t g a a g a g g a t c t c a g g c g t a t a t g a g t c a a c c g a t g a g a a c g a g a c c a a	480
	a c g g a c g t c a g c t t t c g g c a t a t a a c c g a g a t a a c c a t a c t c a c g g t c c a g t t g a t t g t t	540
25	g a g t t t g c t a a a g g t c t a c c a g c g t t t a c a a g a t a c c c c a g g a g g a c c a g a t c a c g t t a	600
	c t a a a g g c c t g c t c g t c g g a g g t g a t g a t g c t g c g t a t g g c a c g a c g c t a t g a c c a c a g c	660
	t c g g a c t c a a t a t t c t t c g c g a a t a t a g a t c a t a t a c g c g g g a t t c t t a c a a a t g g c c	720
	g g a a t g g c t g a t a a c a t t g a a g a c c t g c t g c a t t t c t g c c g c c a a a t g t t c t c g a t g a a g	780
30	g t g g a c a a c g t c g a a t a c g c g t t c t c a c t g c c a t t g t g a t c t t c t c g g a c c g c c g g g c	840
	c t g g a g a a g g c c c a a c t a g t c g a a g c g a t c c a g a g c t a c t a c a t c g a c a c g c t a c g c a t t	900
	t a t a t a c t c a a c c g c c a c t g c g g c g a c t c a a t g a g c c t c t t c t c a c g c a a a g c t g t c t c	960
35	t c g a t c c t c a c c g a g c t g c g t a c g t g g g c a a c c a g a a c g c g a g a t g t t t t c t c a c t a	1020
	a a g c t c a a a a c c g c a a a c t g c c c a a g t t c c t c g a g g a g a t c t g g g a c g t c a t g c c a t c	1080
	c c g c c a t c g g t c c a g t c g c a c c t t c a g a t t a c c c a g g a g g a g a a c g a g c g t c t c g a g c g g	1140
40	g c t g a c g t a t g c g g g c a t c g g t t g g g g g c g c c a t t a c c g c g g c a t t g a t t g c a c t c t	1200
	g c c t c c a c t t c g g c g g c g g c a g c c g c g g c c a g a t c a g c c t a g c c t c a g c c c a g c c c	1260
	c a a c c t c t c t c c t g a c c c a g a a c g a t t c c a g c a c c a g a c a g c c g c a g c t a c a a c t	1320
45	c a g c t a c c a c c t c a g c t g c a a g g t c a a c t g c a a c c c a c g t c c a a c c a c a g c t t c a g a c g	1380

EP 1 266 015 B1

caactccagc cacagattca accacagcca cagctccttc cegttccgc tcccggtccc 1440  
 gcctccgtaa ccgcacctgg ttctttgtcc gcggtcagta cgagcagcga atacatgggc 1500  
 5 ggaagtgcgg ccataggacc catcacgccg gcaaccacca gcagtatcac ggctgcggtt 1560  
 accgctagct ccaccacatc agcgggtacc atgggcaacg gagttggagt cggtgttggg 1620  
 gtggggcgga acgtcagcat gtatgcgaac gccacagcgg cgatggcctt gatgggtgta 1680  
 10 gccctgcatt cgcaccaaga gcagcttata gggggagtgg cggttaagtc ggagcactcg 1740  
 acgactgcac ag 1752

15 <210> 8  
 <211> 1650  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <221> misc feature  
 20 <223> Novel Sequence  
 <400> 8

cgcccggaat gcgtcgtccc ggagaaccaa tgtgcgatga agcggcgaga aaagaaggcc 60  
 25 cagaaggaga aggacaaaat gaccacttcg ccgagctctc agcatggcgg caatggcagc 120  
 ttggcctctg gtggcgccca agactttgtt aagaaggaga ttcttgacct tatgacatgc 180  
 gagcgcgcc agcatgccac tattccgcta ctacctgatg aaatattggc caagtgtcaa 240  
 30 gcgcgaata taccttctct aacgtacaat cagttggcgg ttatatataa gttaatattgg 300  
 taccaggatg gctatgagca gccatctgaa gaggatctca ggcgtataat gagtcaaccc 360  
 gatgagaacg agagccaaac ggacgtcagc ttctggcata taaccgagat aaccatactc 420  
 35 acggtccagt tgattgttga gtttgctaaa ggtctaccag cgtttacaaa gatacccag 480  
 gaggaccaga tcacgttact aaaggcctgc tcgtcggagg tgatgatgct gcgtatggca 540  
 cgacgctatg accacagctc ggaactcaata ttcttcgcga ataatagatc atatacgcgg 600  
 40 gattcttaca aaatggcggg aatggctgat aacattgaag acctgctgca ttctgcgcgc 660  
 caaatgttct cgatgaaggt ggacaacgtc gaatacgcgc ttctcactgc cattgtgata 720  
 ttctcggacc ggcggggcct ggagaaggcc caactagtcg aagcgatcca gagctactac 780  
 45 atcgacacgc tacgcattta tatactcaac cgccactcgc gcgactcaat gagcctcgtc 840  
 ttctacgcaa agctgctctc gatctctacc gagctgcgta cgctgggcaa ccagaacgcc 900  
 gagatgtgtt tctcactaaa gctcaaaaac cgcaaatgc ccaagttcct cgaggagatc 960  
 50 tgggacgttc atgccatccc gccatcggtc cagtcgcacc ttcagattac ccaggaggag 1020  
 aacgagcgtc tcgagcgggc tgagcgtatg cgggcatcgg ttggggggcg cattaccgcc 1080  
 ggcattgatt gcgactctgc ctccacttcg gcggcggcag ccgcggccca gcactcagcct 1140  
 55 cagcctcagc ccagcgcaca acctctctcc ctgaccacaga acgattccca gcaccagaca 1200

	cagccgcagc tacaacctca gctaccacct cagctgcaag gtcaactgca accccagctc	1260
	caaccacagc ttcagacgca actccagcca cagattcaac cacagccaca gctccttccc	1320
5	gtctccgctc ccgtgccgcg ctccgtaacc gcacctgggt ccttgctccg ggctcagtacg	1380
	agcagcgaat acatgggcgg aagtgcggcc ataggaccca tcacgcccgc aaccaccagc	1440
	agtatcacgg ctgccgttac cgctagctcc accacatcag cgttaccgat gggaacaggga	1500
10	gttgagtgctg gtgttgggtg gggcggaac gtacgcatgt atgcgaacgc ccagacggcg	1560
	atggccttga tgggtgtagc cctgcattcg caccaagagc agcttatcgg gggagtggcg	1620
	gttaagtccg agcactcgac gactgcatag	1650
15	<210> 9	
	<211> 1338	
	<212> DNA	
	<213> Artificial Sequence	
20	<220>	
	<221> misc feature	
	<223> Novel Sequence	
	<400> 9	
25	tatgagcagc catctgaaga ggatctcagg cgtataatga gtcaaccgga tgagaacgag	60
	agccaaacgg acgtcagctt tcggcatata accgagataa ccatactcac gggtccagttg	120
	attgttgagt ttgctaaaag tctaccagcg tttaacaaaga taccacagga ggaccagatc	180
30	acgttactaa aggcctgctc gtcggagggtg atgatgctgc gtatggcagc acgctatgac	240
	cacagctcgg actcaatatt ctctcggaat aatagatcat atacgcggga ttcttacaaa	300
	atggccggaa tggctgataa cattgaagac ctgctgcatt tctgccgcca aatgttctcg	360
35	atgaaggctg acaacgtcga atacgcgctt ctactgccca ttgtgatctt ctcggaaccgg	420
	ccgggctcgg agaaggccca actagtcgaa gcgataccaga gctaactcat cgacacgcta	480
	cgcatattata tactcaacgg ccactgcggc gactcaatga gctcgtctt ctacgcaaag	540
40	ctgctctcga tcctcaccca gctgcgtacg ctgggcaacc agaacgcgga gatgtgttct	600
	tcactaaagc tcaaaaacgg caaactgccc aagttcctcg agggagatctg ggacgttcat	660
	gccatcccg ccatcggtcca gtcgcacctt cagattaccc agggaggaga cgagcgtctc	720
45	gagcgggctg agcgtatgcg ggcacgggtt gggggcgcca ttaccgccgg cattgattgc	780
	gactctgcct ccacttcggc ggcggcagcc gcggcccagc atacgcctca gcctcagccc	840
	cagccccaac cctcctcctt gaccagaac gattccagc accagacaca gccgcagcta	900
50	caacctcagc taccacctca gctgcaaggt caactgcaac ccagctcca accacagctt	960
	cagacgcaac tccagccaca gattcaacca cagccacagc tccttcccgt ctccgctccc	1020
	gtgccgcct ccgtaaccgc acctggttcc ttgtccgcgg tcagctacgag cagcgaatac	1080

55

# EP 1 266 015 B1

atggggcgaa gtgcggccat aggacccatc acgcccggcaa ccaccagcag tatcacggct 1140  
 gccgttacgg ctagtccac cacatcagcg gtaccgatgg gcaacggagt tggagtcggg 1200  
 5 gtgggggtgg gcggcaacgt cagcatgtat gcgaacgccc agacggcgat ggccctgatg 1260  
 ggtgtagccc tgcttcgcga ccaagagcag cttatcgggg gagtggcggt taagtcggag 1320  
 cactcgacga ctgcatag 1338

10 <210> 10  
 <211> 969  
 <212> DNA  
 <213> Artificial Sequence  
 15 <220>  
 <221> misc feature  
 <223> Novel Sequence  
 <400> 10

20 cgccgggaat gcgtcgtccc ggagaaccaa tgtgcgatga agcgcgcgga aaagaaggcc 60  
 cagaaggaga aggacaaaaa gaccacttcg ccgagctctc agcatggcgg caatggcagc 120  
 ttggcctctg gtggcgggca agactttgtt aagaaggaga ttcttgacct tatgacatgc 180  
 25 gagccgcccc agcatgccac tattccgcta ctacctgatg aaatattggc caagtgtcaa 240  
 gcgcgaata taccttcctt aacgtacaat cagttggcgg ttatatacaa gttaatttgg 300  
 taccaggatg gctatgagca gccatctgaa gaggatctca ggcgtataat gagtcaaccc 360  
 30 gatgagaacg agagccaaac ggacgtcagc ttccggcata taaccgagat aaccatactc 420  
 acggtccagt tgattgttga gtttgctaaa ggtctaccag cgtttacaaa gataccccgag 480  
 gaggaccaga tcacgttact aaaggcctgc tcgtcggagg tgatgatgct gcgtatggca 540  
 35 cgacgctatg accacagctc ggactcaata ttctcgcga ataatagatc atatacgcg 600  
 gattcttaca aaatggcggg aatggctgat aacattgaag acctgctgca ttctcgccgc 660  
 caaatgttct cgatgaaggt ggacaacgct gaatacgcgc ttctcactgc cattgtgatc 720  
 40 ttctcggacc ggcggggcct ggagaaggcc caactagtcg aagcgatcca gagctactac 780  
 atcgacacgc tacgcattta tatactcaac cgccactgcg gcgactcaat gagcctcgtc 840  
 ttctacgcaa agctgctctc gatcctcacg gagctgcgta cgctggggcaa ccagaacgcc 900  
 45 gagatgtgtt tctcactaaa gctcaaaaac cgcaactgc ccaagttcct cgaggagatc 960  
 tgggacgtt 969

50 <210> 11  
 <211> 412  
 <212> PRT  
 <213> Artificial Sequence  
 <220>  
 <221> misc feature  
 <223> Novel Sequence  
 55 <400> 11

Lys Gly Pro Ala Pro Arg Gln Gln Glu Glu Leu Cys Leu Val Cys Gly  
 1 5 10 15  
 5 Asp Arg Ala Ser Gly Tyr His Tyr Asn Ala Leu Thr Cys Glu Gly Cys  
 20 30  
 Lys Gly Phe Arg Arg Ser Val Thr Lys Asn Ala Val Tyr Ile Cys  
 35 40 45  
 10 Lys Phe Gly His Ala Cys Glu Met Asp Met Tyr Met Arg Arg Lys Cys  
 50 55 60  
 Gln Glu Cys Arg Leu Lys Lys Cys Leu Ala Val Gly Met Arg Pro Glu  
 65 70 75 80  
 15 Cys Val Val Pro Glu Thr Gln Cys Ala Met Lys Arg Lys Glu Lys Lys  
 85 90 95  
 Ala Gln Lys Glu Lys Asp Lys Leu Pro Val Ser Thr Thr Thr Val Asp  
 100 105 110  
 20 Asp His Met Pro Pro Ile Met Gln Cys Glu Pro Pro Pro Glu Ala  
 115 120 125  
 Ala Arg Ile His Glu Val Val Pro Arg Phe Leu Ser Asp Lys Leu Leu  
 130 135 140  
 25 Glu Thr Asn Arg Gln Lys Asn Ile Pro Gln Leu Thr Ala Asn Gln Gln  
 145 150 155 160  
 Phe Leu Ile Ala Arg Leu Ile Trp Tyr Gln Asp Gly Tyr Glu Gln Pro  
 165 170 175  
 30 Ser Asp Glu Asp Leu Lys Arg Ile Thr Gln Thr Trp Gln Gln Ala Asp  
 180 185 190  
 Asp Glu Asn Glu Glu Ser Asp Thr Pro Phe Arg Gln Ile Thr Glu Met  
 195 200 205  
 Thr Ile Leu Thr Val Gln Leu Ile Val Glu Phe Ala Lys Gly Leu Pro  
 210 215 220  
 35 Gly Phe Ala Lys Ile Ser Gln Pro Asp Gln Ile Thr Leu Leu Lys Ala  
 225 230 235 240  
 Cys Ser Ser Glu Val Met Met Leu Arg Val Ala Arg Arg Tyr Asp Ala  
 245 250 255  
 40 Ala Ser Asp Ser Val Leu Phe Ala Asn Asn Gln Ala Tyr Thr Arg Asp  
 260 265 270  
 Asn Tyr Arg Lys Ala Gly Met Ala Tyr Val Ile Glu Asp Leu Leu His  
 275 280 285  
 45 Phe Cys Arg Cys Met Tyr Ser Met Ala Leu Asp Asn Ile His Tyr Ala  
 290 295 300  
 Leu Leu Thr Ala Val Val Ile Phe Ser Asp Arg Pro Gly Leu Glu Gln  
 305 310 315 320  
 50 Pro Gln Leu Val Glu Glu Ile Gln Arg Tyr Tyr Leu Asn Thr Leu Arg

55

EP 1 266 015 B1

325 330 335

Ile Tyr Ile Leu Asn Gln Leu Ser Gly Ser Ala Arg Ser Ser Val Ile  
340 345 350

5 Tyr Gly Lys Ile Leu Ser Ile Leu Ser Glu Leu Arg Thr Leu Gly Met  
355 360 365

Gln Asn Ser Asn Met Cys Ile Ser Leu Lys Leu Lys Asn Arg Lys Leu  
370 375 380

10 Pro Pro Phe Leu Glu Glu Ile Trp Asp Val Ala Asp Met Ser His Thr  
385 390 395 400

Gln Pro Pro Pro Ile Leu Glu Ser Pro Thr Asn Leu  
405 410

15 <210> 12  
<211> 412  
<212> PRT  
<213> Artificial Sequence  
20 <220>  
<221> misc\_feature  
<223> Novel Sequence  
<400> 12

25 Lys Gly Pro Ala Pro Arg Gln Gln Glu Glu Leu Cys Leu Val Cys Gly  
1 5 10 15

Asp Arg Ala Ser Gly Tyr His Tyr Asn Ala Leu Thr Cys Glu Gly Cys  
20 25 30

30 Lys Gly Phe Phe Arg Arg Ser Val Thr Lys Asn Ala Val Tyr Ile Cys  
35 40 45

Lys Phe Gly His Ala Cys Glu Met Asp Met Tyr Met Arg Arg Lys Cys  
50 55 60

35 Gln Glu Cys Arg Leu Lys Lys Cys Leu Ala Val Gly Met Arg Pro Glu  
65 70 75 80

Cys Val Val Pro Glu Thr Gln Cys Ala Met Lys Arg Lys Glu Lys Lys  
85 90 95

40 Ala Gln Lys Glu Lys Asp Lys Leu Pro Val Ser Thr Thr Thr Val Asp  
100 105 110

Asp His Met Pro Pro Ile Met Gln Cys Glu Pro Pro Pro Pro Glu Ala  
115 120 125

45 Ala Arg Ile His Glu Val Val Pro Arg Phe Leu Ser Asp Lys Leu Leu  
130 135 140

Glu Thr Asn Arg Gln Lys Asn Ile Pro Gln Leu Thr Ala Asn Gln Gln  
145 150 155 160

50 Phe Leu Ile Ala Arg Leu Ile Trp Tyr Gln Asp Gly Tyr Glu Gln Pro  
165 170 175

Ser Asp Glu Asp Leu Lys Arg Ile Thr Gln Thr Trp Gln Gln Ala Asp  
180 185 190

55

EP 1 266 015 B1

Asp Glu Asn Glu Glu Ser Asp Thr Pro Phe Arg Gln Ile Thr Glu Met  
 195 200 205  
 Thr Ile Leu Thr Val Gln Leu Ile Val Glu Phe Ala Lys Gly Leu Pro  
 210 215 220  
 Gly Phe Ala Lys Ile Ser Gln Pro Asp Gln Ile Thr Leu Leu Lys Ala  
 225 230 235 240  
 Cys Ser Ser Glu Val Met Met Leu Arg Val Ala Arg Arg Tyr Asp Ala  
 245 250 255  
 Ala Ser Asp Ser Val Leu Phe Ala Asn Asn Gln Ala Tyr Thr Arg Asp  
 260 265 270  
 Asn Tyr Arg Lys Ala Gly Met Ala Tyr Val Ile Glu Asp Leu Leu His  
 275 280 285  
 Phe Cys Arg Cys Met Tyr Ser Met Ala Leu Asp Asn Ile His Tyr Ala  
 290 295 300  
 Leu Leu Thr Ala Val Val Ile Phe Ser Asp Arg Pro Gly Leu Glu Gln  
 305 310 315 320  
 Pro Gln Leu Val Glu Glu Ile Gln Arg Tyr Tyr Leu Asn Thr Leu Arg  
 325 330 335  
 Ile Tyr Ile Leu Asn Gln Leu Ser Gly Ser Ala Arg Ser Ser Val Ile  
 340 345 350  
 Tyr Gly Lys Ile Leu Ser Ile Leu Ser Glu Leu Arg Thr Leu Gly Met  
 355 360 365  
 Gln Asn Ser Asn Met Cys Ile Ser Leu Lys Leu Lys Asn Arg Lys Leu  
 370 375 380  
 Pro Pro Phe Leu Glu Glu Ile Trp Asp Val Ala Asp Met Ser His Thr  
 385 390 395 400  
 Gln Pro Pro Pro Ile Leu Glu Ser Pro Thr Asn Leu  
 405 410

<210> 13  
 <211> 334  
 <212> PRT  
 <213> Artificial Sequence  
 <220>  
 <221> misc feature  
 <223> Novel Sequence  
 <400> 13

Pro Glu Cys Val Val Pro Glu Thr Gln Cys Ala Met Lys Arg Lys Glu  
 1 5 10 15  
 Lys Lys Ala Gln Lys Glu Lys Asp Lys Leu Pro Val Ser Thr Thr Thr  
 20 25 30  
 Val Asp Asp His Met Pro Pro Ile Met Gln Cys Glu Pro Pro Pro Pro  
 35 40 45  
 Glu Ala Ala Arg Ile His Glu Val Val Pro Arg Phe Leu Ser Asp Lys  
 50 55 60

5 Leu Leu Glu Thr Asn Arg Gln Lys Asn Ile Pro Gln Leu Thr Ala Asn  
 65 70 75 80  
 Gln Gln Phe Leu Ile Ala Arg Leu Ile Trp Tyr Gln Asp Gly Tyr Glu  
 85 90 95  
 Gln Pro Ser Asp Glu Asp Leu Lys Arg Ile Thr Gln Thr Trp Gln Gln  
 100 105 110  
 10 Ala Asp Asp Glu Asn Glu Glu Ser Asp Thr Pro Phe Arg Gln Ile Thr  
 115 120 125  
 Glu Met Thr Ile Leu Thr Val Gln Leu Ile Val Glu Phe Ala Lys Gly  
 130 135 140  
 15 Leu Pro Gly Phe Ala Lys Ile Ser Gln Pro Asp Gln Ile Thr Leu Leu  
 145 150 155 160  
 Lys Ala Cys Ser Ser Glu Val Met Met Leu Arg Val Ala Arg Arg Tyr  
 165 170 175  
 20 Asp Ala Ala Ser Asp Ser Val Leu Phe Ala Asn Asn Gln Ala Tyr Thr  
 180 185 190  
 Arg Asp Asn Tyr Arg Lys Ala Gly Met Ala Tyr Val Ile Glu Asp Leu  
 195 200 205  
 25 Leu His Phe Cys Arg Cys Met Tyr Ser Met Ala Leu Asp Asn Ile His  
 210 215 220  
 Tyr Ala Leu Leu Thr Ala Val Val Ile Phe Ser Asp Arg Pro Gly Leu  
 225 230 235 240  
 Glu Gln Pro Gln Leu Val Glu Glu Ile Gln Arg Tyr Tyr Leu Asn Thr  
 245 250 255  
 30 Leu Arg Ile Tyr Ile Leu Asn Gln Leu Ser Gly Ser Ala Arg Ser Ser  
 260 265 270  
 Val Ile Tyr Gly Lys Ile Leu Ser Ile Leu Ser Glu Leu Arg Thr Leu  
 275 280 285  
 35 Gly Met Gln Asn Ser Asn Met Cys Ile Ser Leu Lys Leu Lys Asn Arg  
 290 295 300  
 Lys Leu Pro Pro Phe Leu Glu Glu Ile Trp Asp Val Ala Asp Met Ser  
 305 310 315 320  
 40 His Thr Gln Pro Pro Pro Ile Leu Glu Ser Pro Thr Asn Leu  
 325 330

<210> 14  
 <211> 244  
 <212> PRT  
 <213> Artificial Sequence  
 <220>  
 <221> misc feature  
 <223> Novel Sequence  
 <400> 14

Tyr Gln Asp Gly Tyr Glu Gln Pro Ser Asp Glu Asp Leu Lys Arg Ile



EP 1 266 015 B1

	1		5		10		15									
	Thr	Gln	Thr	Trp 20	Gln	Gln	Ala	Asp	Asp 25	Glu	Asn	Glu	Glu	Ser	Asp	Thr
5	Pro	Phe	Arg 35	Gln	Ile	Thr	Glu	Met 40	Thr	Ile	Leu	Thr	Val 45	Gln	Leu	Ile
	Val	Glu	Phe	Ala	Lys	Gly	Leu 55	Pro	Gly	Phe	Ala	Lys 60	Ile	Ser	Gln	Pro
10	Asp 65	Gln	Ile	Thr	Leu 70	Lys	Ala	Cys	Ser 75	Glu	Val	Met	Met	Leu 80		
	Arg	Val	Ala	Arg	Arg 85	Tyr	Asp	Ala	Ala 90	Ser	Asp	Ser	Val	Leu	Phe	Ala
15	Asn	Asn	Gln	Ala 100	Tyr	Thr	Arg	Asp	Asn 105	Tyr	Arg	Lys	Ala	Gly 110	Met	Ala
	Tyr	Val	Ile 115	Glu	Asp	Leu	Leu	His 120	Phe	Cys	Arg	Cys	Met 125	Tyr	Ser	Met
20	Ala	Leu	Asp	Asn	Ile	His	Tyr 135	Ala	Leu	Leu	Thr	Ala 140	Val	Val	Ile	Phe
	Ser	Asp	Arg	Pro	Gly	Leu 150	Glu	Gln	Pro	Gln	Leu 155	Val	Glu	Glu	Ile	Gln
25	Arg	Tyr	Tyr	Leu 165	Asn	Thr	Leu	Arg	Ile 170	Tyr	Ile	Leu	Asn	Gln	Leu	Ser
	Gly	Ser	Ala	Arg 180	Ser	Ser	Val	Ile	Tyr 185	Gly	Lys	Ile	Leu	Ser 190	Ile	Leu
30	Ser	Glu	Leu 195	Arg	Thr	Leu	Gly	Met 200	Gln	Asn	Ser	Asn	Met 205	Cys	Ile	Ser
	Leu	Lys	Leu	Lys	Asn	Arg	Lys 215	Leu	Pro	Pro	Phe	Leu 220	Glu	Glu	Ile	Trp
35	Asp 225	Val	Ala	Asp	Met	Ser 230	His	Thr	Gln	Pro	Pro 235	Pro	Ile	Leu	Glu	Ser 240
	Pro	Thr	Asn	Leu												
40	<210> 15 <211> 320 <212> PRT <213> Artificial Sequence <220>															
45	<221> misc feature <223> Novel Sequence <400> 15															
50	Pro	Glu	Cys	Val 5	Val	Pro	Glu	Thr	Gln 10	Cys	Ala	Met	Lys	Arg	Lys	Glu 15
	Lys	Lys	Ala	Gln 20	Lys	Glu	Lys	Asp	Lys 25	Leu	Pro	Val	Ser	Thr 30	Thr	Thr
55																

EP 1 266 015 B1

Val Asp Asp His Met Pro Pro Ile Met Gln Cys Glu Pro Pro Pro Pro  
35 40 45

5 Glu Ala Ala Arg Ile His Glu Val Val Pro Arg Phe Leu Ser Ser Asp Lys  
50 55 60

Leu Leu Glu Thr Asn Arg Gln Lys Asn Ile Pro Gln Leu Thr Ala Asn  
65 70 80

10 Gln Gln Phe Leu Ile Ala Arg Leu Ile Trp Tyr Gln Asp Gly Tyr Glu  
85 90 95

Gln Pro Ser Asp Glu Asp Leu Lys Arg Ile Thr Gln Thr Trp Gln Gln  
100 105 110

15 Ala Asp Asp Glu Asn Glu Glu Ser Asp Thr Pro Phe Arg Gln Ile Thr  
115 120 125

Glu Met Thr Ile Leu Thr Val Gln Leu Ile Val Glu Phe Ala Lys Gly  
130 135 140

20 Leu Pro Gly Phe Ala Lys Ile Ser Gln Pro Asp Gln Ile Thr Leu Leu  
145 150 155 160

Lys Ala Cys Ser Ser Glu Val Met Met Leu Arg Val Ala Arg Arg Tyr  
165 170 175

25 Asp Ala Ala Ser Asp Ser Val Leu Phe Ala Asn Asn Gln Ala Tyr Thr  
180 185 190

Arg Asp Asn Tyr Arg Lys Ala Gly Met Ala Tyr Val Ile Glu Asp Leu  
195 200 205

30 Leu His Phe Cys Arg Cys Met Tyr Ser Met Ala Leu Asp Asn Ile His  
210 215 220

Tyr Ala Leu Leu Thr Ala Val Val Ile Phe Ser Asp Arg Pro Gly Leu  
225 230 235 240

35 Glu Gln Pro Gln Leu Val Glu Glu Ile Gln Arg Tyr Tyr Leu Asn Thr  
245 250 255

Leu Arg Ile Tyr Ile Leu Asn Gln Leu Ser Gly Ser Ala Arg Ser Ser  
260 265 270

40 Val Ile Tyr Gly Lys Ile Leu Ser Ile Leu Ser Glu Leu Arg Thr Leu  
275 280 285

Gly Met Gln Asn Ser Asn Met Cys Ile Ser Leu Lys Leu Lys Asn Arg  
290 295 300

Lys Leu Pro Pro Phe Leu Glu Glu Ile Trp Asp Val Ala Asp Met Ser  
305 310 315 320

<210> 16  
<211> 625  
<212> PRT  
<213> Artificial Sequence  
50 <220>  
<221> misc feature  
<223> Novel Sequence  
<400> 16

Gly Pro Ala Pro Arg Val Gln Glu Glu Leu Cys Leu Val Cys Gly Asp  
 1 5 10 15  
 Arg Ala Ser Gly Tyr His Tyr Asn Ala Leu Thr Cys Glu Gly Cys Lys  
 20 30  
 Gly Phe Phe Arg Arg Ser Val Thr Lys Ser Ala Val Tyr Cys Cys Lys  
 35 40 45  
 Phe Gly Arg Ala Cys Glu Met Asp Met Tyr Met Arg Lys Cys Gln  
 50 55 60  
 Glu Cys Arg Leu Lys Lys Cys Leu Ala Val Gly Met Arg Pro Glu Cys  
 65 70 75 80  
 Val Val Pro Glu Asn Gln Cys Ala Met Lys Arg Arg Glu Lys Lys Ala  
 85 90 95  
 Gln Lys Glu Lys Asp Lys Met Thr Thr Ser Pro Ser Ser Gln His Gly  
 100 105 110  
 Gly Asn Gly Ser Leu Ala Ser Gly Gly Gln Asp Phe Val Lys Lys  
 115 120 125  
 Glu Ile Leu Asp Leu Met Thr Cys Glu Pro Pro Gln His Ala Thr Ile  
 130 135 140  
 Pro Leu Leu Pro Asp Glu Ile Leu Ala Lys Cys Gln Ala Arg Asn Ile  
 145 150 155 160  
 Pro Ser Leu Thr Tyr Asn Gln Leu Ala Val Ile Tyr Lys Leu Ile Trp  
 165 170 175  
 Tyr Gln Asp Gly Tyr Glu Gln Pro Ser Glu Glu Asp Leu Arg Arg Ile  
 180 185 190  
 Met Ser Gln Pro Asp Glu Asn Glu Ser Gln Thr Asp Val Ser Phe Arg  
 195 200 205  
 His Ile Thr Glu Ile Thr Ile Leu Thr Val Gln Leu Ile Val Glu Phe  
 210 215 220  
 Ala Lys Gly Leu Pro Ala Phe Thr Lys Ile Pro Gln Glu Asp Gln Ile  
 225 230 235 240  
 Thr Leu Leu Lys Ala Cys Ser Ser Glu Val Met Met Leu Arg Met Ala  
 245 250 255  
 Arg Arg Tyr Asp His Ser Ser Asp Ser Ile Phe Phe Ala Asn Asn Arg  
 260 265 270  
 Ser Tyr Thr Arg Asp Ser Tyr Lys Met Ala Gly Met Ala Asp Asn Ile  
 275 280 285  
 Glu Asp Leu Leu His Phe Cys Arg Gln Met Phe Ser Met Lys Val Asp  
 290 295 300  
 Asn Val Glu Tyr Ala Leu Leu Thr Ala Ile Val Ile Phe Ser Asp Arg  
 305 310 315 320  
 Pro Gly Leu Glu Lys Ala Gln Leu Val Glu Ala Ile Gln Ser Tyr Tyr  
 325 330 335  
 Ile Asp Thr Leu Arg Ile Tyr Ile Leu Asn Arg His Cys Gly Asp Ser

EP 1 266 015 B1

	340	345	350
	Met Ser Leu Val Phe Tyr Ala Lys Leu Leu Ser Ile Leu Thr Glu Leu		
	355	360	365
5	Arg Thr Leu Gly Asn Gln Asn Ala Glu Met Cys Phe Ser Leu Lys Leu		
	370	375	380
	Lys Asn Arg Lys Leu Pro Lys Phe Leu Glu Glu Ile Trp Asp Val His		
	385	390	400
10	Ala Ile Pro Pro Ser Val Gln Ser His Leu Gln Ile Thr Gln Glu Glu		
	405	410	415
	Asn Glu Arg Leu Glu Arg Ala Glu Arg Met Arg Ala Ser Val Gly Gly		
	420	425	430
15	Ala Ile Thr Ala Gly Ile Asp Cys Asp Ser Ala Ser Thr Ser Ala Ala		
	435	440	445
	Ala Ala Ala Ala Gln His Gln Pro Gln Pro Gln Pro Gln Pro Gln Pro		
	450	455	460
20	Ser Ser Leu Thr Gln Asn Asp Ser Gln His Gln Thr Gln Pro Gln Leu		
	465	470	475
	Gln Pro Gln Leu Pro Pro Gln Leu Gln Gly Gln Leu Gln Pro Gln Leu		
	485	490	495
25	Gln Pro Gln Leu Gln Thr Gln Leu Gln Pro Gln Ile Gln Pro Gln Pro		
	500	505	510
	Gln Leu Leu Pro Val Ser Ala Pro Val Pro Ala Ser Val Thr Ala Pro		
	515	520	525
30	Gly Ser Leu Ser Ala Val Ser Thr Ser Ser Glu Tyr Met Gly Gly Ser		
	530	535	540
	Ala Ala Ile Gly Pro Ile Thr Pro Ala Thr Thr Ser Ser Ile Thr Ala		
	545	550	555
35	Ala Val Thr Ala Ser Ser Thr Thr Ser Ala Val Pro Met Gly Asn Gly		
	565	570	575
	Val Gly Val Gly Val Gly Val Gly Asn Val Ser Met Tyr Ala Asn		
	580	585	590
40	Ala Gln Thr Ala Met Ala Leu Met Gly Val Ala Leu His Ser His Gln		
	595	600	605
	Glu Gln Leu Ile Gly Gly Val Ala Val Lys Ser Glu His Ser Thr Thr		
	610	615	620
45	Ala		
	625		

<210> 17  
 <211> 583  
 <212> PRT  
 <213> Artificial Sequence  
 <220>  
 <221> misc feature  
 <223> Novel Sequence  
 <400> 17

Ala Val Tyr Cys Cys Lys Phe Gly Arg Ala Cys Glu Met Asp Met Tyr  
 1 5 10 15  
 Met Arg Arg Lys Cys Gln Glu Cys Arg Leu Lys Lys Cys Leu Ala Val  
 20 25 30  
 Gly Met Arg Pro Glu Cys Val Val Pro Glu Asn Gln Cys Ala Met Lys  
 35 40 45  
 Arg Arg Glu Lys Lys Ala Gln Lys Glu Lys Asp Lys Met Thr Thr Ser  
 50 55 60  
 Pro Ser Ser Gln His Gly Gly Asn Gly Ser Leu Ala Ser Gly Gly Gly  
 65 70 75 80  
 Gln Asp Phe Val Lys Lys Glu Ile Leu Asp Leu Met Thr Cys Glu Pro  
 85 90 95  
 Pro Gln His Ala Thr Ile Pro Leu Leu Pro Asp Glu Ile Leu Ala Lys  
 100 105 110  
 Cys Gln Ala Arg Asn Ile Pro Ser Leu Thr Tyr Asn Gln Leu Ala Val  
 115 120 125  
 Ile Tyr Lys Leu Ile Trp Tyr Gln Asp Gly Tyr Glu Gln Pro Ser Glu  
 130 135 140  
 Glu Asp Leu Arg Arg Ile Met Ser Gln Pro Asp Glu Asn Glu Ser Gln  
 145 150 155 160  
 Thr Asp Val Ser Phe Arg His Ile Thr Glu Ile Thr Ile Leu Thr Val  
 165 170 175  
 Gln Leu Ile Val Glu Phe Ala Lys Gly Leu Pro Ala Phe Thr Lys Ile  
 180 185 190  
 Pro Gln Glu Asp Gln Ile Thr Leu Leu Lys Ala Cys Ser Ser Glu Val  
 195 200 205  
 Met Met Leu Arg Met Ala Arg Arg Tyr Asp His Ser Ser Asp Ser Ile  
 210 215 220  
 Phe Phe Ala Asn Asn Arg Ser Tyr Thr Arg Asp Ser Tyr Lys Met Ala  
 225 230 235 240  
 Gly Met Ala Asp Asn Ile Glu Asp Leu Leu His Phe Cys Arg Gln Met  
 245 250 255  
 Phe Ser Met Lys Val Asp Asn Val Glu Tyr Ala Leu Leu Thr Ala Ile  
 260 265 270  
 Val Ile Phe Ser Asp Arg Pro Gly Leu Glu Lys Ala Gln Leu Val Glu  
 275 280 285  
 Ala Ile Gln Ser Tyr Tyr Ile Asp Thr Leu Arg Ile Tyr Ile Leu Asn  
 290 295 300  
 Arg His Cys Gly Asp Ser Met Ser Leu Val Phe Tyr Ala Lys Leu Leu  
 305 310 315 320  
 Ser Ile Leu Thr Glu Leu Arg Thr Leu Gly Asn Gln Asn Ala Glu Met  
 325 330 335

EP 1 266 015 B1

Cys Phe Ser Leu Lys Leu Lys Asn Arg Lys Leu Pro Lys Phe Leu Glu  
 340 345 350  
 5 Glu Ile Trp Asp Val His Ala Ile Pro Pro Ser Val Gln Ser His Leu  
 355 360 365  
 Gln Ile Thr Gln Glu Glu Asn Glu Arg Leu Glu Arg Ala Glu Arg Met  
 370 375 380  
 10 Arg Ala Ser Val Gly Gly Ala Ile Thr Ala Gly Ile Asp Cys Asp Ser  
 385 390 395 400  
 Ala Ser Thr Ser Ala Ala Ala Ala Ala Gln His Gln Pro Gln Pro  
 405 410 415  
 15 Gln Pro Gln Pro Gln Pro Ser Ser Leu Thr Gln Asn Asp Ser Gln His  
 420 425 430  
 Gln Thr Gln Pro Gln Leu Gln Pro Gln Leu Pro Pro Gln Leu Gln Gly  
 435 440 445  
 20 Gln Leu Gln Pro Gln Leu Gln Pro Gln Leu Gln Thr Gln Leu Gln Pro  
 450 455 460  
 Gln Ile Gln Pro Gln Pro Gln Leu Leu Pro Val Ser Ala Pro Val Pro  
 465 470 475 480  
 25 Ala Ser Val Thr Ala Pro Gly Ser Leu Ser Ala Val Ser Thr Ser Ser  
 485 490 495  
 Glu Tyr Met Gly Gly Ser Ala Ala Ile Gly Pro Ile Thr Pro Ala Thr  
 500 505 510  
 30 Thr Ser Ser Ile Thr Ala Ala Val Thr Ala Ser Ser Thr Thr Ser Ala  
 515 520 525  
 Val Pro Met Gly Asn Gly Val Gly Val Gly Val Gly Val Gly Asn  
 530 535 540  
 Val Ser Met Tyr Ala Asn Ala Gln Thr Ala Met Ala Leu Met Gly Val  
 545 550 555 560  
 35 Ala Leu His Ser His Gln Glu Gln Leu Ile Gly Gly Val Ala Val Lys  
 565 570 575  
 Ser Glu His Ser Thr Thr Ala  
 580  
 40 <210> 18  
 <211> 549  
 <212> PRT  
 <213> Artificial Sequence  
 45 <400> 18  
 Arg Pro Glu Cys Val Val Pro Glu Asn Gln Cys Ala Met Lys Arg Arg  
 1 5 10 15  
 50 Glu Lys Lys Ala Gln Lys Glu Lys Asp Lys Met Thr Thr Ser Pro Ser  
 20 25 30  
 Ser Gln His Gly Gly Asn Gly Ser Leu Ala Ser Gly Gly Gly Gln Asp  
 35 40 45  
 55

EP 1 266 015 B1

Phe Val Lys Lys Glu Ile Leu Asp Leu Met Thr Cys Glu Pro Pro Gln  
 50 55 60  
 5 His Ala Thr Ile Pro Leu Leu Pro Asp Glu Ile Leu Ala Lys Cys Gln  
 65 70 75 80  
 Ala Arg Asn Ile Pro Ser Leu Thr Tyr Asn Gln Leu Ala Val Ile Tyr  
 85 90 95  
 10 Lys Leu Ile Trp Tyr Gln Asp Gly Tyr Glu Gln Pro Ser Glu Glu Asp  
 100 105 110  
 Leu Arg Arg Ile Met Ser Gln Pro Asp Glu Asn Glu Ser Gln Thr Asp  
 115 120 125  
 15 Val Ser Phe Arg His Ile Thr Glu Ile Thr Ile Leu Thr Val Gln Leu  
 130 135 140  
 Ile Val Glu Phe Ala Lys Gly Leu Pro Ala Phe Thr Lys Ile Pro Gln  
 145 150 155 160  
 20 Glu Asp Gln Ile Thr Leu Leu Lys Ala Cys Ser Ser Glu Val Met Met  
 165 170 175  
 Leu Arg Met Ala Arg Arg Tyr Asp His Ser Ser Asp Ser Ile Phe Phe  
 180 185 190  
 25 Ala Asn Asn Arg Ser Tyr Thr Arg Asp Ser Tyr Lys Met Ala Gly Met  
 195 200 205  
 Ala Asp Asn Ile Glu Asp Leu Leu His Phe Cys Arg Gln Met Phe Ser  
 210 215 220  
 30 Met Lys Val Asp Asn Val Glu Tyr Ala Leu Leu Thr Ala Ile Val Ile  
 225 230 235 240  
 Phe Ser Asp Arg Pro Gly Leu Glu Lys Ala Gln Leu Val Glu Ala Ile  
 245 250 255  
 Gln Ser Tyr Tyr Ile Asp Thr Leu Arg Ile Tyr Ile Leu Asn Arg His  
 260 265 270  
 35 Cys Gly Asp Ser Met Ser Leu Val Phe Tyr Ala Lys Leu Ser Ile  
 275 280 285  
 Leu Thr Glu Leu Arg Thr Leu Gly Asn Gln Asn Ala Glu Met Cys Phe  
 290 295 300  
 40 Ser Leu Lys Leu Lys Asn Arg Lys Leu Pro Lys Phe Leu Glu Glu Ile  
 305 310 315 320  
 Trp Asp Val His Ala Ile Pro Pro Ser Val Gln Ser His Leu Gln Ile  
 325 330 335  
 45 Thr Gln Glu Glu Asn Glu Arg Leu Glu Arg Ala Glu Arg Met Arg Ala  
 340 345 350  
 Ser Val Gly Gly Ala Ile Thr Ala Gly Ile Asp Cys Asp Ser Ala Ser  
 355 360 365  
 50 Thr Ser Ala Ala Ala Ala Ala Ala Gln His Gln Pro Gln Pro Gln Pro  
 370 375 380  
 Gln Pro Gln Pro Ser Ser Leu Thr Gln Asn Asp Ser Gln His Gln Thr  
 385 390 395 400

EP 1 266 015 B1

5 Gln Pro Gln Leu Gln Pro Gln Leu Pro Pro Gln Leu Gln Gly Gln Leu  
 405 410 415  
 Gln Pro Gln Leu Gln Pro Gln Leu Gln Thr Gln Leu Gln Pro Gln Ile  
 420 425 430  
 Gln Pro Gln Pro Gln Leu Leu Pro Val Ser Ala Pro Val Pro Ala Ser  
 435 440 445  
 10 Val Thr Ala Pro Gly Ser Leu Ser Ala Val Ser Thr Ser Ser Glu Tyr  
 450 455 460  
 Met Gly Gly Ser Ala Ala Ile Gly Pro Ile Thr Pro Ala Thr Thr Ser  
 465 470 475 480  
 15 Ser Ile Thr Ala Ala Val Thr Ala Ser Ser Thr Thr Ser Ala Val Pro  
 485 490 495  
 Met Gly Asn Gly Val Gly Val Gly Val Gly Val Gly Gly Asn Val Ser  
 500 505 510  
 20 Met Tyr Ala Asn Ala Gln Thr Ala Met Ala Leu Met Gly Val Ala Leu  
 515 520 525  
 His Ser His Gln Glu Gln Leu Ile Gly Gly Val Ala Val Lys Ser Glu  
 530 535 540  
 His Ser Thr Thr Ala  
 545  
 25  
 <210> 19  
 <211> 445  
 <212> PRT  
 30 <213> Artificial Sequence  
 <400> 19  
 35 Tyr Glu Gln Pro Ser Glu Glu Asp Leu Arg Arg Ile Met Ser Gln Pro  
 1 5 10 15  
 Asp Glu Asn Glu Ser Gln Thr Asp Val Ser Phe Arg His Ile Thr Glu  
 20 25 30  
 40 Ile Thr Ile Leu Thr Val Gln Leu Ile Val Glu Phe Ala Lys Gly Leu  
 35 40 45  
 Pro Ala Phe Thr Lys Ile Pro Gln Glu Asp Gln Ile Thr Leu Leu Lys  
 50 55 60  
 45 Ala Cys Ser Ser Glu Val Met Met Leu Arg Met Ala Arg Arg Tyr Asp  
 65 70 75 80  
 His Ser Ser Asp Ser Ile Phe Phe Ala Asn Asn Arg Ser Tyr Thr Arg  
 85 90 95  
 Asp Ser Tyr Lys Met Ala Gly Met Ala Asp Asn Ile Glu Asp Leu Leu  
 100 105 110  
 50 His Phe Cys Arg Gln Met Phe Ser Met Lys Val Asp Asn Val Glu Tyr  
 115 120 125  
 Ala Leu Leu Thr Ala Ile Val Ile Phe Ser Asp Arg Pro Gly Leu Glu  
 130 135 140  
 55



EP 1 266 015 B1

5           Lys Ala Gln Leu Val Glu Ala Ile Gln Ser Tyr Tyr Ile Asp Thr Leu  
           145                   150                   155                   160  
 Arg Ile Tyr Ile Leu Asn Arg His Cys Gly Asp Ser Met Ser Leu Val  
           165                   170                   175  
 Phe Tyr Ala Lys Leu Leu Ser Ile Leu Thr Glu Leu Arg Thr Leu Gly  
           180                   185                   190  
 10        Asn Gln Asn Ala Glu Met Cys Phe Ser Leu Lys Leu Lys Asn Arg Lys  
           195                   200                   205  
 Leu Pro Lys Phe Leu Glu Glu Ile Trp Asp Val His Ala Ile Pro Pro  
           210                   215                   220  
 15        Ser Val Gln Ser His Leu Gln Ile Thr Gln Glu Glu Asn Glu Arg Leu  
           225                   230                   235                   240  
 Glu Arg Ala Glu Arg Met Arg Ala Ser Val Gly Gly Ala Ile Thr Ala  
           245                   250                   255  
 20        Gly Ile Asp Cys Asp Ser Ala Ser Thr Ser Ala Ala Ala Ala Ala Ala  
           260                   265                   270  
 Gln His Gln Pro Gln Pro Gln Pro Gln Pro Ser Ser Leu Thr  
           275                   280                   285  
 25        Gln Asn Asp Ser Gln His Gln Thr Gln Pro Gln Leu Gln Pro Gln Leu  
           290                   295                   300  
 Pro Pro Gln Leu Gln Gly Gln Leu Gln Pro Gln Leu Gln Pro Gln Leu  
           305                   310                   315                   320  
 Gln Thr Gln Leu Gln Pro Gln Ile Gln Pro Gln Pro Gln Leu Leu Pro  
           325                   330                   335  
 30        Val Ser Ala Pro Val Pro Ala Ser Val Thr Ala Pro Gly Ser Leu Ser  
           340                   345                   350  
 Ala Val Ser Thr Ser Ser Glu Tyr Met Gly Gly Ser Ala Ala Ile Gly  
           355                   360                   365  
 35        Pro Ile Thr Pro Ala Thr Thr Ser Ser Ile Thr Ala Ala Val Thr Ala  
           370                   375                   380  
 Ser Ser Thr Thr Ser Ala Val Pro Met Gly Asn Gly Val Gly Val Gly  
           385                   390                   395                   400  
 40        Val Gly Val Gly Gly Asn Val Ser Met Tyr Ala Asn Ala Gln Thr Ala  
           405                   410                   415  
 Met Ala Leu Met Gly Val Ala Leu His Ser His Gln Glu Gln Leu Ile  
           420                   425                   430  
 45        Gly Gly Val Ala Val Lys Ser Glu His Ser Thr Thr Ala  
           435                   440                   445

<210> 20

<211> 323

<212> PRT

<213> Artificial Sequence

<400> 20

55           Arg Pro Glu Cys Val Val Pro Glu Asn Gln Cys Ala Met Lys Arg Arg

EP 1 266 015 B1

	1	5	10	15
5	Glu Lys Lys	Ala Gln Lys Glu Lys	Asp Lys Met Thr Thr	Ser Pro Ser
		20	25	30
	Ser Gln His	Gly Gly Asn Gly Ser	Leu Ala Ser Gly Gly	Gly Gln Asp
	35	40	45	
10	Phe Val Lys	Lys Glu Ile Leu Asp	Leu Met Thr Cys Glu	Pro Pro Gln
	50	55	60	
	His Ala Thr	Ile Pro Leu Leu Pro	Asp Glu Ile Leu Ala	Lys Cys Gln
	65	70	75	80
15	Ala Arg Asn	Ile Pro Ser Leu Thr	Tyr Asn Gln Leu Ala	Val Ile Tyr
		85	90	95
	Lys Leu Ile	Trp Tyr Gln Asp Gly	Tyr Glu Gln Pro Ser	Glu Glu Asp
		100	105	110
20	Leu Arg Arg	Ile Met Ser Gln Pro	Asp Glu Asn Glu Ser	Gln Thr Asp
		115	120	125
	Val Ser Phe	Arg His Ile Thr Glu	Ile Thr Ile Leu Thr	Val Gln Leu
	130	135	140	
	Ile Val Glu	Phe Ala Lys Gly Leu	Pro Ala Phe Thr Lys	Ile Pro Gln
	145	150	155	160
25	Glu Asp Gln	Ile Thr Leu Leu Lys	Ala Cys Ser Ser Glu	Val Met Met
		165	170	175
	Leu Arg Met	Ala Arg Arg Tyr Asp	His Ser Ser Asp Ser	Ile Phe Phe
		180	185	190
30	Ala Asn Asn	Arg Ser Tyr Thr Arg	Asp Ser Tyr Lys Met	Ala Gly Met
		195	200	205
	Ala Asp Asn	Ile Glu Asp Leu Leu	His Phe Cys Arg Gln	Met Phe Ser
		210	215	220
35	Met Lys Val	Asp Asn Val Glu Tyr	Ala Leu Leu Thr Ala	Ile Val Ile
		225	230	235
	Phe Ser Asp	Arg Pro Gly Leu Glu	Lys Ala Gln Leu Val	Glu Ala Ile
		245	250	255
40	Gln Ser Tyr	Tyr Ile Asp Thr Leu	Arg Ile Tyr Ile Leu	Asn Arg His
		260	265	270
	Cys Gly Asp	Ser Met Ser Leu Val	Phe Tyr Ala Lys Leu	Leu Ser Ile
		275	280	285
45	Leu Thr Glu	Leu Arg Thr Leu Gly	Asn Gln Asn Ala Glu	Met Cys Phe
		290	295	300
	Ser Leu Lys	Leu Lys Asn Arg Lys	Leu Pro Lys Phe Leu	Glu Glu Ile
		305	310	315
50	Trp Asp Val			320

<210> 21

<211> 987

<212> DNA

55 <213> Artificial Sequence

<400> 21

tgtgctatct gtggggaccg ctcctcagcg aaacactatg gggatatacag ttgtgagggc 60  
 tgcaagggct tcttcaagag gacagtacgc aaagacctga cctacacctg ccgagacaaac 120  
 5 aaggactgcc tgatcgacaa gagacagcgg aaccggtgtc agtactgccg ctaccagaag 180  
 tgcctggcca tgggcatgaa gcgggaaagct gtgcaggagg agcggcagcg gggcaaggac 240  
 cggaaatgaga acgaggtgga gtccaccagc agtgccaacg aggacatgcc tgtagagaag 300  
 10 attctggaag ccgagcttgc tgctcgagccc aagactgaga catacgtgga ggcaaacatg 360  
 gggctgaacc ccagctcacc aaatgaccct gttaccaaca tctgtcaacg agcagacaaag 420  
 cagctcttca ctcttgtgga gtgggccaag aggateccac acttttctga gctgccctta 480  
 15 gacgaccagg tcactctgct acgggcagcg tggaaacgagc tgctgatcgc ctcttctctc 540  
 caccgttcca tagctgtgaa agatgggatt ctctggcca ccggcctgca cgtacaccgg 600  
 aacagcgctc acagtgtctg ggtgggcgcc atctttgaca ggtgtctaac agagctggtg 660  
 20 tctaagatgc gtgacatgca gatggacaag acggagctgg gctgcctgcg agccaattgtc 720  
 ctgttcaacc ctgactctaa ggggctctca aacctgctg aggtggaggc gttgagggag 780  
 aagggtgatg cgtcactaga agcgtactgc aaacacaagt accctgagca gccgggcagg 840  
 25 ttgccaagc tgctgctcgg cctgcctgca ctgcgttcca tcgggctcaa gtgcctggag 900  
 cactgttct tcttcaagct catcggggac acgcccatcg acacttcct catggagatg 960  
 ctggaggcac cacatcaagc cacctag 987

30 <210> 22  
 <211> 789  
 <212> DNA  
 <213> Artificial Sequence  
 <400> 22

35 aagcgggaag ctgtgcagga ggagcggcag cggggcaagg accggaatga gaacagggg 60  
 gagtccacca gcagtcccaa cgaggacatg cctgtagaga agattctgga agccgagctt 120  
 40 gctgtcgagc ccaagactga gacatacgtg gagggcaaca tggggctgaa cccagctca 180  
 ccaaatgacc ctgttaccaa catctgtcaa gcagcagaca agcagctctt cactcttgtg 240  
 gagtgggcca agaggatccc acacttttct gagctgcccc tagacgacca ggtcatcctg 300  
 45 ctacgggcag gctggaacga gctgctgacg gcctcttctt cccaccgctc catagctgtg 360  
 aaagatggga ttctcctgpc caccggcctg cactacacc ggaacagcgc tcacagtgtc 420  
 ggggtgggag ccatctttga cagggtgcta acagagctgg tgtctaagat gcgtgacatg 480  
 50 cagatggaca agacggagct gggctgcctg cgagccattg tctgttcaa cctgactct 540  
 aaggggctct caaacctcgc tgaggtggag gcgttgaggg agaaggtgta tgcgtcacta 600  
 gaagcgtact gcaaacacaa gtaccctgag cagccgggca ggtttgcca catgctgctc 660

55

cgccctgcctg cactgcgttc catcggggctc aagtgcctgg agcacctgtt cttcttcaag 720  
 ctcacatgggg acacgcccac cgacaccttc ctcacggaga tgcctggaggc accacatcaa 780  
 gccacctag 789

<210> 23  
 <211> 714  
 <212> DNA  
 <213> Artificial Sequence  
 <400> 23

gccaacgagg acatgcctgt agagaagatt ctggaagccg agcttctgt cgagcccaag 60  
 actgagacat acgtggaggc aaacatgggg ctgaacccca gctcaccaaa tgacctgtt 120  
 accaacatct gtcaagcagc agacaagcag ctcttcactc ttgtggagtg ggccaagagg 180  
 atcccacact tttctgagct gccccctagac gaccagggtca tctctgtacg ggcaggctgg 240  
 aacgagctgc tgatgcctc cttctccacc cgctccatag ctgtgaaaga tgggattctc 300  
 ctggccaccg gctgcacgt acaccggaac agcgctcaca gtgctgggtg gggcgccatc 360  
 tttagacagg tgctaacaga gctggtgtct aagatgcgtg acatgcagat ggacaagacg 420  
 gagctgggct gctgcgagc cattgtcctg ttcacccctg actctaaggc gctctcaaac 480  
 cctgctgagg tggaggcgtt gagggagaag gtgtatgcgt cactagaagg ctactgcaaa 540  
 cacaagtacc ctgagcagcc gggcagggtt gccaaagctgc tgctccgcct gcctgcactg 600  
 cgttccatcg ggctcaagtg cctggagcac ctgttcttct tcaagctcat cggggacacg 660  
 cccatcgaca ctttctcat ggagatgctg gaggcaccac atcaagccac ctg 714

<210> 24  
 <211> 536  
 <212> DNA  
 <213> Artificial Sequence  
 <400> 24

ggatcccaca cttttctgag ctgcccctag acgaccagggt catctgtcta cgggcaggct 60  
 ggaacgagct gctgatcgcc tcttctctcc accgctccat agctgtgaaa gatgggattc 120  
 tcctggccac cggcctgcac gtacaccgga acagcgctca cagtgtctggg gtgggcgccca 180  
 tctttgacag ggtgctaaca gagctggtgt ctaagatgcg tgacatgcag atggacaaga 240  
 cggagctggg ctgcctgcga gccattgtcc tgttcaaccc tgactctaag gggctctcaa 300  
 accctgctga ggtggaggcg ttgagggaga aggtgtatgc gtcaactagaa gcgtactgca 360  
 aacacaagta ccctgagcag ccggggcaggt ttgccaagct gctgctccgc ctgcctgcac 420  
 tgcgttccat cgggctcaag tgcctggagc acctgttctt cttcaagctc atcggggaca 480  
 cgcccatcga cacttctctc atggagatgc tggaggcacc acatcaagcc acctag 536

<210> 25  
 <211> 672

# EP 1 266 015 B1

<212> DNA

<213> Artificial Sequence

<400> 25

5	gccaacgagg acatgcctgt agagaagatt ctggaagccg agcttgctgt cgagcccaag	60
	actgagacat acgtggaggc aaacatgggg ctgaacccca gctcaccaaa tgacctgttt	120
	accaacatct gtcaagcagc agacaagcag ctcttcactc ttgtggagtg ggcgaagagg	180
10	atccacact tttctgagct gcccttagac gaccagggtc tctgtctacg ggcaggctgg	240
	aaagagctgc tgatgcctc cttctccac cgctccatag ctgtgaaaga tgggattctc	300
	ctggccaccg gcctgcacgt acaccggaac agcgctcaca gtgctggggt ggcgccatc	360
15	tttgacagg tgtaacaga gctggtgtct aagatgctg acatgcagat ggacaagacg	420
	gagctgggct gcctgcgagc cattgtcctg ttcaaccctg actctaaggg gctctcaaac	480
	cctgctgagg tggaggcgtt gagggagaag gtgtatgctg cactagaagc gtactgcaaa	540
20	cacaagtacc ctgagcagcc gggcagggtt gccaaagctg tgctccgcct gcctgcactg	600
	cgttccatcg ggctcaagtg cctggagcac ctgttcttct tcaagctcat cggggacacg	660
	cccatcgaca cc	672

<210> 26

<211> 1123

<212> DNA

<213> Artificial Sequence

<400> 26

30	tgcccatct gcggggaccg ctctcagcg aagcactatg gagtgtacag ctgcgagggg	60
	tgcaagggct tcttcaagcg gacggtgcgc aaggacctga cctacacctg ccgcgacaac	120
35	aaggactgcc tgattgacaa gcggcagcgg aaccgggtgcc agtactgccg ctaccagaag	180
	tgccctggcca tgggcatgaa gcgggaagcc gtgcaggagg agcggcagcg tggcaaggac	240
	cggaacgaga atgagggtgga gtcgaccagc agcgccaacg aggacatgcc ggtggagagg	300
40	atcctggagg ctgagctggc cgtggagccc aagaccgaga cctacgtgga ggcaaacatg	360
	gggctgaacc ccagctgcgc gaacgacctc gtcaccaaca tttgccaaag agccgacaaa	420
	cagcttttca cctcgttgga gtgggccaag cggatcccac acttctcaga gctgccctg	480
45	gacgaccagg tcactcgtct gcgggcagcg tggaatgagc tgctcatcgc ctcttctcc	540
	caccgctcca tcgccgtgaa ggacgggata ctcttgccca ccgggctgca cgtccaccgg	600
	aacagcgccc acagcgaggg ggtgggcgac atctttgaca ggggtgctgac ggagcttggt	660
50	tccaagatgc gggacatgca gatggacaag acggagctgg gctgcctgag cgccatcgct	720
	ctctttaacc ctgactccaa ggggctctcg aaccggccg aggtggaggc gctgaggagg	780
	aagggtctatg cgtccttgga ggcctactgc aagcacaagt acccagagca gccgggaagg	840
55	ttcgctaagc tcttgctccg cctgcggcgt ctgcgtcca tcgggctcaa atgcctggaa	900

catctcttct tcttcaagct catcggggac acaccattg acaccttct tatggagatg 960  
 ctggaggcgc cgcaccaaat gacttaggcc tgcgggcccc tcctttgtgc ccaaccgttc 1020  
 tggccacctt gcttggagcg cagctgttct tctcagcctg agccctgtcc ctgccttctt 1080  
 ctgcctggcc tgtttggact ttggggcaca gctgtctact gct 1123

<210> 27  
 <211> 925  
 <212> DNA  
 <213> Artificial Sequence  
 <400> 27

aagcgggaag ccgtgcagga ggagcgccag cgtggcaagg accggaacga gaatgaggtg 60  
 gaggcgacca gcagcgccaa cgaggacatg ccggtggaga ggatcctgga ggctgagctg 120  
 gccgtggagc ccaagaccga gacctacgtg gaggcaaaaca tggggctgaa cccagctcg 180  
 ccgaacgacc ctgtcaccaa catttgccaa gcagccgaca aacagctttt caccctggtg 240  
 gagtgggcca agcggatccc acacttctca gacttgcccc tggagaccca ggctatcctg 300  
 ctgcgggcaa gctggaatga gctgctcact gcctccttct cccaccgctc catcgccgtg 360  
 aaggacggga tcctcctggc caccgggctg cactgcccac ggaacagcgc ccacagcgca 420  
 ggggtgggcg ccatctttga cagggtgctg acggagcttg tgtccaaagt gcgggacatg 480  
 cagatggaca agacggagct gggctgacct gcgcgccatg tcctctttaa cctgactcc 540  
 aaggggctct cgaaccgagc cgagggtggag gcgctgaggg agaaggtcta tgcgtccttg 600  
 gaggcctact gcaagcaca gtaccagag cagccgggaa ggttcgctaa gctcttgctc 660  
 cgctgcggg ctctgcgctc catcgggctc aaatgcctgg aacatctctt ctcttcaag 720  
 ctcatcgggg acacaccat tgacaccttc cttatggaga tgcgtgaggg gccgcaccaa 780  
 atgacttagg cctgcgggac catcctttgt gcccaaccgt tctggccacc ctgcctggac 840  
 gccagctgtt ctctcagcgc tgagccctgt cctgcccctt ctctgcctgg cctgtttgga 900  
 ctttggggca cagcctgtca ctgct 925

<210> 28  
 <211> 850  
 <212> DNA  
 <213> Artificial Sequence  
 <400> 28

gccaacgagg acatgcgggt ggagaggatc ctggaggctg agctggccgt ggagcccaag 60  
 accgagacct acgtggaggc aaacatgggg ctgaacccca gctcgcgaa cgaccctgtc 120  
 accaacattt gccaaagcgc cgacaaacag cttttcaccc tggtggaagt ggccaagcgg 180  
 atccacactt tctcagagct gccctggac gaccaggcca tcctgctgct ggagagctgg 240  
 aatgagctgc tcactgcctc ctctctccac gcctccatcg ccgtgaagga cgggaccttc 300

ctggccaccg ggctgcacgt ccaccggaac agcgcccaca gcgcaggggg gggcgccatc 360  
 tttgacaggg tgctgacgga gcttgtgtcc aagatgcggg acatgcagat ggacaagacg 420  
 5 gactggggct gcctgcgcgc catcgtcttc ttaaccctg actccaagg gctctcgaac 480  
 ccggccgagg tggaggcgtt gagggagaag gtctatgcgt ccttgagggc ctactgcaag 540  
 cacaagtacc cagagcagcc gggaagggtc gctaagctct tgctccgcct gccggtcttg 600  
 10 cgctccatcg ggctcaaatg cctggaacat ctcttcttct tcaagctcat cggggacaca 660  
 cccattgaca ccttccttat ggagatgctg gaggcgccgc accaaatgac ttaggcctgc 720  
 gggcccatcc tttgtgccca ccggttcttg ccaccctgcc tggacgccag ctgtttcttc 780  
 15 cagcctgagc cctgtccctg cccttctctg cctggcctgt ttggaacttg gggcacagcc 840  
 tgtcactgct 850

<210> 29  
 <211> 670  
 <212> DNA  
 <213> Artificial Sequence  
 <400> 29

25 atcccacact tctcagagct gcccctggac gaccagggtca tctctgctgc ggcaggctgg 60  
 aatgagctgc tcctgccttc cttctccac cgctccatcg ccgtgaaggc cggatccttc 120  
 ctggccaccg ggctgcacgt ccaccggaac agcgcccaca gcgcaggggg gggcgccatc 180  
 30 tttgacaggg tgctgacgga gcttgtgtcc aagatgcggg acatgcagat ggacaagacg 240  
 gactggggct gcctgcgcgc catcgtcttc ttaaccctg actccaagg gctctcgaac 300  
 ccggccgagg tggaggcgtt gagggagaag gtctatgcgt ccttgagggc ctactgcaag 360  
 35 cacaagtacc cagagcagcc gggaagggtc gctaagctct tgctccgcct gccggtcttg 420  
 cgctccatcg ggctcaaatg cctggaacat ctcttcttct tcaagctcat cggggacaca 480  
 cccattgaca ccttccttat ggagatgctg gaggcgccgc accaaatgac ttaggcctgc 540  
 40 gggcccatcc tttgtgccca ccggttcttg ccaccctgcc tggacgccag ctgtttcttc 600  
 cagcctgagc cctgtccctg cccttctctg cctggcctgt ttggaacttg gggcacagcc 660  
 tgtcactgct 670

45 <210> 30  
 <211> 672  
 <212> DNA  
 <213> Artificial Sequence  
 50 <400> 30

gccaacgagg acatgccggt ggagaggatc ctggaggctg agctggccgt ggagccaag 60  
 accgagacct acgtggaggc aaacatgggg ctgaacccca gctcgccgaa cgaccctgtc 120  
 55 accaaccatt gccaaagcgc cgacaacacg cttttcacc ttggtgagtg ggccaagcgg 180  
 atcccacact tctcagagct gcccctggac gaccagggtca tctctgctgc ggcaggctgg 240

aatgagctgc tcatgcctc cttctccac cgtccatcg cgtgaagga cgggatcctc 300  
 ctggccacgg ggtgcacgt ccaccggaac agcgcacaca gcgcaggggt gggcgccatc 360  
 5 tttgacaggg tgctgacgga gcttggtgcc aagatgcggg acatgcagat ggacaagacg 420  
 gagctggggt gcttgcgcg catcgtcctc tttaccctg actccaagg gctctcgaac 480  
 cggccgaggg tggaggcgct gagggagaag gtctatgct ccttgaggc ctactgcaag 540  
 10 cacaagtacc cagagcagcc gggaaggtc gctaagctct tgctcggct gccggctctg 600  
 cgctccatcg ggctcaaat cctggaacat ctctctctct tcaagctcat cggggacaca 660  
 cccattgaca cc 672

<210> 31  
 <211> 328  
 <212> PRT  
 <213> Artificial Sequence  
 20 <400> 31

Cys Ala Ile Cys Gly Asp Arg Ser Ser Gly Lys His Tyr Gly Val Tyr  
 1 5 10 15  
 25 Ser Cys Glu Gly Cys Lys Gly Phe Phe Lys Arg Thr Val Arg Lys Asp  
 20 25 30  
 Leu Thr Tyr Thr Cys Arg Asp Asn Lys Asp Cys Leu Ile Asp Lys Arg  
 35 35 40 45  
 30 Gln Arg Asn Arg Cys Gln Tyr Cys Arg Tyr Gln Lys Cys Leu Ala Met  
 50 55 60  
 Gly Met Lys Arg Glu Ala Val Gln Glu Glu Arg Gln Arg Gly Lys Asp  
 65 70 75 80  
 35 Arg Asn Glu Asn Glu Val Glu Ser Thr Ser Ala Asn Glu Asp Met  
 85 90 95  
 Pro Val Glu Lys Ile Leu Glu Ala Glu Leu Ala Val Glu Pro Lys Thr  
 100 105 110  
 40 Glu Thr Tyr Val Glu Ala Asn Met Gly Leu Asn Pro Ser Ser Pro Asn  
 115 120 125  
 Asp Pro Val Thr Asn Ile Cys Gln Ala Ala Asp Lys Gln Leu Phe Thr  
 130 135 140  
 45 Leu Val Glu Trp Ala Lys Arg Ile Pro His Phe Ser Glu Leu Pro Leu  
 145 150 155 160  
 Asp Asp Gln Val Ile Leu Leu Arg Ala Gly Trp Asn Glu Leu Leu Ile  
 165 170 175  
 50 Ala Ser Phe Ser His Arg Ser Ile Ala Val Lys Asp Gly Ile Leu Leu  
 180 185 190  
 Ala Thr Gly Leu His Val His Arg Asn Ser Ala His Ser Ala Gly Val  
 195 200 205  
 55 Gly Ala Ile Phe Asp Arg Val Leu Thr Glu Leu Val Ser Lys Met Arg



EP 1 266 015 B1

210 215 220

Asp Met Gln Met Asp Lys Thr Glu Leu Gly Cys Leu Arg Ala Ile Val  
225 230 235 240

Leu Phe Asn Pro Asp Ser Lys Gly Leu Ser Asn Pro Ala Glu Val Glu  
245 250 255

Ala Leu Arg Glu Lys Val Tyr Ala Ser Leu Glu Ala Tyr Cys Lys His  
260 265 270

Lys Tyr Pro Glu Gln Pro Gly Arg Phe Ala Lys Leu Leu Arg Leu  
275 280 285

Pro Ala Leu Arg Ser Ile Gly Leu Lys Cys Leu Glu His Leu Phe Phe  
290 295 300

Phe Lys Leu Ile Gly Asp Thr Pro Ile Asp Thr Phe Leu Met Glu Met  
305 310 315 320

Leu Glu Ala Pro His Gln Ala Thr  
325

<210> 32  
<211> 262  
<212> PRT  
<213> Artificial Sequence  
<400> 32

Lys Arg Glu Ala Val Gln Glu Glu Arg Gln Arg Gly Lys Asp Arg Asn  
1 5 10 15

Glu Asn Glu Val Glu Ser Thr Ser Ser Ala Asn Glu Asp Met Pro Val  
20 25 30

Glu Lys Ile Leu Glu Ala Glu Leu Ala Val Glu Pro Lys Thr Glu Thr  
35 40 45

Tyr Val Glu Ala Asn Met Gly Leu Asn Pro Ser Ser Pro Asn Asp Pro  
50 55 60

Val Thr Asn Ile Cys Gln Ala Ala Asp Lys Gln Leu Phe Thr Leu Val  
65 70 75 80

Glu Trp Ala Lys Arg Ile Pro His Phe Ser Glu Leu Pro Leu Asp Asp  
85 90 95

Gln Val Ile Leu Leu Arg Ala Gly Trp Asn Glu Leu Leu Ile Ala Ser  
100 105 110

Phe Ser His Arg Ser Ile Ala Val Lys Asp Gly Ile Leu Leu Ala Thr  
115 120 125

Gly Leu His Val His Arg Asn Ser Ala His Ser Ala Gly Val Gly Ala  
130 135 140

Ile Phe Asp Arg Val Leu Thr Glu Leu Val Ser Lys Met Arg Asp Met  
145 150 155 160

Gln Met Asp Lys Thr Glu Leu Gly Cys Leu Arg Ala Ile Val Leu Phe  
165 170 175

Asn Pro Asp Ser Lys Gly Leu Ser Asn Pro Ala Glu Val Glu Ala Leu  
180 185 190

Arg Glu Lys Val Tyr Ala Ser Leu Glu Ala Tyr Cys Lys His Lys Tyr  
 195 200 205  
 5 Pro Glu Gln Pro Gly Arg Phe Ala Lys Leu Leu Leu Arg Leu Pro Ala  
 210 215 220  
 Leu Arg Ser Ile Gly Leu Lys Cys Leu Glu His Leu Phe Phe Phe Lys  
 225 230 235 240  
 10 Leu Ile Gly Asp Thr Pro Ile Asp Thr Phe Leu Met Glu Met Leu Glu  
 245 250 255  
 Ala Pro His Gln Ala Thr  
 260  
 15 <210> 33  
 <211> 237  
 <212> PRT  
 <213> Artificial Sequence  
 <400> 33  
 20  
 Ala Asn Glu Asp Met Pro Val Glu Lys Ile Leu Glu Ala Glu Leu Ala  
 1 5 10 15  
 25 Val Glu Pro Lys Thr Glu Thr Tyr Val Glu Ala Asn Met Gly Leu Asn  
 20 25 30  
 Pro Ser Ser Pro Asn Asp Pro Val Thr Asn Ile Cys Gln Ala Ala Asp  
 35 40 45  
 Lys Gln Leu Phe Thr Leu Val Glu Trp Ala Lys Arg Ile Pro His Phe  
 50 55 60  
 30 Ser Glu Leu Pro Leu Asp Gln Val Ile Leu Leu Arg Ala Gly Trp  
 65 70 75 80  
 Asn Glu Leu Leu Ile Ala Ser Phe Ser His Arg Ser Ile Ala Val Lys  
 85 90 95  
 35 Asp Gly Ile Leu Leu Ala Thr Gly Leu His Val His Arg Asn Ser Ala  
 100 105 110  
 His Ser Ala Gly Val Gly Ala Ile Phe Asp Arg Val Leu Thr Glu Leu  
 115 120 125  
 40 Val Ser Lys Met Arg Asp Met Gln Met Asp Lys Thr Glu Leu Gly Cys  
 130 135 140  
 Leu Arg Ala Ile Val Leu Phe Asn Pro Asp Ser Lys Gly Leu Ser Asn  
 145 150 155 160  
 45 Pro Ala Glu Val Glu Ala Leu Arg Glu Lys Val Tyr Ala Ser Leu Glu  
 165 170 175  
 Ala Tyr Cys Lys His Lys Tyr Pro Glu Gln Pro Gly Arg Phe Ala Lys  
 180 185 190  
 50 Leu Leu Leu Arg Leu Pro Ala Leu Arg Ser Ile Gly Leu Lys Cys Leu  
 195 200 205  
 Glu His Leu Phe Phe Phe Lys Leu Ile Gly Asp Thr Pro Ile Asp Thr  
 210 215 220  
 55

EP 1 266 015 B1

Phe Leu Met Glu Met Leu Glu Ala Pro His Gln Ala Thr  
225 230 235

<210> 34  
<211> 177  
<212> PRT  
<213> Artificial Sequence  
<400> 34

Ile Pro His Phe Ser Glu Leu Pro Leu Asp Asp Gln Val Ile Leu Leu  
1 5 10 15  
Arg Ala Gly Trp Asn Glu Leu Leu Ile Ala Ser Phe Ser His Arg Ser  
20 25 30  
Ile Ala Val Lys Asp Gly Ile Leu Leu Ala Thr Gly Leu His Val His  
35 40 45  
Arg Asn Ser Ala His Ser Ala Gly Val Gly Ala Ile Phe Asp Arg Val  
50 55 60  
Leu Thr Glu Leu Val Ser Lys Met Arg Asp Met Gln Met Asp Lys Thr  
65 70 75 80  
Glu Leu Gly Cys Leu Arg Ala Ile Val Leu Phe Asn Pro Asp Ser Lys  
85 90 95  
Gly Leu Ser Asn Pro Ala Glu Val Glu Ala Leu Arg Glu Lys Val Tyr  
100 105 110  
Ala Ser Leu Glu Ala Tyr Cys Lys His Lys Tyr Pro Glu Gln Pro Gly  
115 120 125  
Arg Phe Ala Lys Leu Leu Leu Arg Leu Pro Ala Leu Arg Ser Ile Gly  
130 135 140  
Leu Lys Cys Leu Glu His Leu Phe Phe Phe Lys Leu Ile Gly Asp Thr  
145 150 155 160  
Pro Ile Asp Thr Phe Leu Met Glu Met Leu Glu Ala Pro His Gln Ala  
165 170 175  
Thr

<210> 35  
<211> 224  
<212> PRT  
<213> Artificial Sequence  
<400> 35

Ala Asn Glu Asp Met Pro Val Glu Lys Ile Leu Glu Ala Glu Leu Ala  
1 5 10 15  
Val Glu Pro Lys Thr Glu Thr Tyr Val Glu Ala Asn Met Gly Leu Asn  
20 25 30  
Pro Ser Ser Pro Asn Asp Pro Val Thr Asn Ile Cys Gln Ala Ala Asp  
35 40 45  
Lys Gln Leu Phe Thr Leu Val Glu Trp Ala Lys Arg Ile Pro His Phe  
50 55 60

EP 1 266 015 B1

Ser Glu Leu Pro Leu Asp Asp Gln Val Ile Leu Leu Arg Ala Gly Trp  
 65 70 75 80  
 Asn Glu Leu Leu Ile Ala Ser Phe Ser His Arg Ser Ile Ala Val Lys  
 85 90 95  
 Asp Gly Ile Leu Leu Ala Thr Gly Leu His Val His Arg Asn Ser Ala  
 100 105 110  
 His Ser Ala Gly Val Gly Ala Ile Phe Asp Arg Val Leu Thr Glu Leu  
 115 120 125  
 Val Ser Lys Met Arg Asp Met Gln Met Asp Lys Thr Glu Leu Gly Cys  
 130 135 140  
 Leu Arg Ala Ile Val Leu Phe Asn Pro Asp Ser Lys Gly Leu Ser Asn  
 145 150 155 160  
 Pro Ala Glu Val Glu Ala Leu Arg Glu Lys Val Tyr Ala Ser Leu Glu  
 165 170 175  
 Ala Tyr Cys Lys His Lys Tyr Pro Glu Gln Pro Gly Arg Phe Ala Lys  
 180 185 190  
 Leu Leu Leu Arg Leu Pro Ala Leu Arg Ser Ile Gly Leu Lys Cys Leu  
 195 200 205  
 Glu His Leu Phe Phe Phe Lys Leu Ile Gly Asp Thr Pro Ile Asp Thr  
 210 215 220

<210> 36  
 <211> 328  
 <212> PRT  
 <213> Artificial Sequence  
 <220>  
 <221> misc feature  
 <223> Novel Sequence  
 <400> 36

Cys Ala Ile Cys Gly Asp Arg Ser Ser Gly Lys His Tyr Gly Val Tyr  
 1 5 10 15  
 Ser Cys Glu Gly Cys Lys Gly Phe Phe Lys Arg Thr Val Arg Lys Asp  
 20 25 30  
 Leu Thr Tyr Thr Cys Arg Asp Asn Lys Asp Cys Leu Ile Asp Lys Arg  
 35 40 45  
 Gln Arg Asn Arg Cys Gln Tyr Cys Arg Tyr Gln Lys Cys Leu Ala Met  
 50 55 60  
 Gly Met Lys Arg Glu Ala Val Gln Glu Glu Arg Gln Arg Gly Lys Asp  
 65 70 75 80  
 Arg Asn Glu Asn Glu Val Glu Ser Thr Ser Ser Ala Asn Glu Asp Met  
 85 90 95  
 Pro Val Glu Arg Ile Leu Glu Ala Glu Leu Ala Val Glu Pro Lys Thr  
 100 105 110  
 Glu Thr Tyr Val Glu Ala Asn Met Gly Leu Asn Pro Ser Ser Pro Asn  
 115 120 125

Asp Pro Val Thr Asn Ile Cys Gln Ala Ala Asp Lys Gln Leu Phe Thr  
 130 135 140  
 5 Leu Val Glu Trp Ala Lys Arg Ile Pro His Phe Ser Glu Leu Pro Leu  
 145 150 155  
 Asp Asp Gln Val Ile Leu Leu Arg Ala Gly Trp Asn Glu Leu Leu Ile  
 165 170 175  
 10 Ala Ser Phe Ser His Arg Ser Ile Ala Val Lys Asp Gly Ile Leu Leu  
 180 185 190  
 Ala Thr Gly Leu His Val His Arg Asn Ser Ala His Ser Ala Gly Val  
 195 200 205  
 15 Gly Ala Ile Phe Asp Arg Val Leu Thr Glu Leu Val Ser Lys Met Arg  
 210 215 220  
 Asp Met Gln Met Asp Lys Thr Glu Leu Gly Cys Leu Arg Ala Ile Val  
 225 230 235 240  
 20 Leu Phe Asn Pro Asp Ser Lys Gly Leu Ser Asn Pro Ala Glu Val Glu  
 245 250 255  
 Ala Leu Arg Glu Lys Val Tyr Ala Ser Leu Glu Ala Tyr Cys Lys His  
 260 265 270  
 25 Lys Tyr Pro Glu Gln Pro Gly Arg Phe Ala Lys Leu Leu Leu Arg Leu  
 275 280 285  
 Pro Ala Leu Arg Ser Ile Gly Leu Lys Cys Leu Glu His Leu Phe Phe  
 290 295 300  
 30 Phe Lys Leu Ile Gly Asp Thr Pro Ile Asp Thr Phe Leu Met Glu Met  
 305 310 315 320  
 Leu Glu Ala Pro His Gln Met Thr  
 325  
 <210> 37  
 35 <211> 262  
 <212> PRT  
 <213> Artificial Sequence  
 <220>  
 <221> misc feature  
 40 <223> Novel Sequence  
 <400> 37  
 Lys Arg Glu Ala Val Gln Glu Glu Arg Gln Arg Gly Lys Asp Arg Asn  
 1 5 10 15  
 Glu Asn Glu Val Glu Ser Thr Ser Ser Ala Asn Glu Asp Met Pro Val  
 20 25 30  
 50 Glu Arg Ile Leu Glu Ala Glu Leu Ala Val Glu Pro Lys Thr Glu Thr  
 35 40 45  
 Tyr Val Glu Ala Asn Met Gly Leu Asn Pro Ser Ser Pro Asn Asp Pro  
 50 55 60  
 Val Thr Asn Ile Cys Gln Ala Ala Asp Lys Gln Leu Phe Thr Leu Val

EP 1 266 015 B1

[illegible]

```

35      <210> 38
        <211> 237
        <212> PRT
        <213> Artificial Sequence
        <220>
        <221> misc feature
        <223> Novel Sequence
40      <400> 38

```

	Ala	Asn	Glu	Asp	Met	Pro	Val	Glu	Arg	Ile	Leu	Glu	Ala	Glu	Leu	Ala
	1				5					10					15	
45	Val	Glu	Pro	Lys	Thr	Glu	Thr	Tyr	Val	Glu	Ala	Asn	Met	Gly	Leu	Asn
				20					25					30		
	Pro	Ser	Ser	Pro	Asn	Asp	Pro	Val	Thr	Asn	Ile	Cys	Gln	Ala	Ala	Asp
			35					40					45			
50	Lys	Gln	Leu	Phe	Thr	Leu	Val	Glu	Trp	Ala	Lys	Arg	Ile	Pro	His	Phe
		50					55					60				
	Ser	Glu	Leu	Pro	Leu	Asp	Gln	Val	Ile	Leu	Leu	Arg	Ala	Gly	Trp	
	65					70				75					80	

EP 1 266 015 B1

Asn Glu Leu Leu Ile Ala Ser Phe Ser His Arg Ser Ile Ala Val Lys  
85 90 95

5 Asp Gly Ile Leu Leu Ala Thr Gly Leu His Val His Arg Asn Ser Ala  
100 105 110

His Ser Ala Gly Val Gly Ala Ile Phe Asp Arg Val Leu Thr Glu Leu  
115 120 125

10 Val Ser Lys Met Arg Asp Met Gln Met Asp Lys Thr Glu Leu Gly Cys  
130 135 140

Leu Arg Ala Ile Val Leu Phe Asn Pro Asp Ser Lys Gly Leu Ser Asn  
145 150 155 160

15 Pro Ala Glu Val Glu Ala Leu Arg Glu Lys Val Tyr Ala Ser Leu Glu  
165 170 175

Ala Tyr Cys Lys His Lys Tyr Pro Glu Gln Pro Gly Arg Phe Ala Lys  
180 185 190

20 Leu Leu Leu Arg Leu Pro Ala Leu Arg Ser Ile Gly Leu Lys Cys Leu  
195 200 205

Glu His Leu Phe Phe Phe Lys Leu Ile Gly Asp Thr Pro Ile Asp Thr  
210 215 220

25 Phe Leu Met Glu Met Leu Glu Ala Pro His Gln Met Thr  
225 230 235

<210> 39  
<211> 177  
<212> PRT  
30 <213> Artificial Sequence  
<220>  
<221> misc\_feature  
<223> Novel Sequence  
35 <400> 39

Ile Pro His Phe Ser Glu Leu Pro Leu Asp Asp Gln Val Ile Leu Leu  
1 5 10 15

40 Arg Ala Gly Trp Asn Glu Leu Leu Ile Ala Ser Phe Ser His Arg Ser  
20 25 30

Ile Ala Val Lys Asp Gly Ile Leu Leu Ala Thr Gly Leu His Val His  
35 40 45

45 Arg Asn Ser Ala His Ser Ala Gly Val Gly Ala Ile Phe Asp Arg Val  
50 55 60

Leu Thr Glu Leu Val Ser Lys Met Arg Asp Met Gln Met Asp Lys Thr  
65 70 75 80

50 Glu Leu Gly Cys Leu Arg Ala Ile Val Leu Phe Asn Pro Asp Ser Lys  
85 90 95

Gly Leu Ser Asn Pro Ala Glu Val Glu Ala Leu Arg Glu Lys Val Tyr  
100 105 110

55 Ala Ser Leu Glu Ala Tyr Cys Lys His Lys Tyr Pro Glu Gln Pro Gly  
115 120 125

Arg Phe Ala Lys Leu Leu Arg Leu Pro Ala Leu Arg Ser Ile Gly  
130 135 140

Leu Lys Cys Leu Glu His Leu Phe Phe Phe Lys Leu Ile Gly Asp Thr  
145 150 155 160

Pro Ile Asp Thr Phe Leu Met Glu Met Leu Glu Ala Pro His Gln Met  
165 170 175

Thr

<210> 40

<211> 224

<212> PRT

<213> Artificial Sequence

<220>

<221> misc feature

<223> Novel Sequence

<400> 40

Ala Asn Glu Asp Met Pro Val Glu Arg Ile Leu Glu Ala Glu Leu Ala  
1 5 10 15

Val Glu Pro Lys Thr Glu Thr Tyr Val Glu Ala Asn Met Gly Leu Asn  
20 25 30

Pro Ser Ser Pro Asn Asp Pro Val Thr Asn Ile Cys Gln Ala Ala Asp  
35 40 45

Lys Gln Leu Phe Thr Leu Val Glu Trp Ala Lys Arg Ile Pro His Phe  
50 55 60

Ser Glu Leu Pro Leu Asp Asp Gln Val Ile Leu Arg Ala Gly Trp  
65 70 75 80

Asn Glu Leu Leu Ile Ala Ser Phe Ser His Arg Ser Ile Ala Val Lys  
85 90 95

Asp Gly Ile Leu Leu Ala Thr Gly Leu His Val His Arg Asn Ser Ala  
100 105 110

His Ser Ala Gly Val Gly Ala Ile Phe Asp Arg Val Leu Thr Glu Leu  
115 120 125

Val Ser Lys Met Arg Asp Met Gln Met Asp Lys Thr Glu Leu Gly Cys  
130 135 140

Leu Arg Ala Ile Val Leu Phe Asn Pro Asp Ser Lys Gly Leu Ser Asn  
145 150 155 160

Pro Ala Glu Val Glu Ala Leu Arg Glu Lys Val Tyr Ala Ser Leu Glu  
165 170 175

Ala Tyr Cys Lys His Lys Tyr Pro Glu Gln Pro Gly Arg Phe Ala Lys  
180 185 190

Leu Leu Leu Arg Leu Pro Ala Leu Arg Ser Ile Gly Leu Lys Cys Leu  
195 200 205

Glu His Leu Phe Phe Phe Lys Leu Ile Gly Asp Thr Pro Ile Asp Thr

210

215

220



<210> 41  
 <211> 441  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <221> misc feature  
 <223> Novel Sequence  
 <400> 41

```

10      atgaagctac tgtcttctat cgaacaagca tgcgatattt gccgacttaa aaagctcaag      60
      tgctccaaag aaaaaccgaa gtgcgccaag tgtctgaaga acaactggga gtgtcgctac      120
      tctcccaaaa caaaagggtc tccgctgact agggcacatc tgacagaagt ggaatcaagg      180
15      ctagaaagac tggaacagct atttctactg attttctctc gagaagacct tgacatgatt      240
      ttgaaaatgg attctttaca ggatataaaa gcattgttaa caggattatt tgtacaagat      300
      aatgtgaata aagatgccgt cacagataga ttggcttcag tggagactga tatgcctcta      360
20      acattgagac agcatagaat aagtgcgaca tcatcatcgg aagagagtag taacaaagggt      420
      caaagacagt tgactgtatc g                                441

```

<210> 42  
 <211> 147  
 <212> PRT  
 <213> Artificial Sequence  
 <220>  
 <221> misc\_feature  
 <223> Novel Sequence  
 <400> 42

```

35      Met Lys Leu Leu Ser Ser Ile Glu Gln Ala Cys Asp Ile Cys Arg Leu
      1          5          10          15
      Lys Lys Leu Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu
      20          25          30
40      Lys Asn Asn Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro
      35          40          45
      Leu Thr Arg Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu
      50          55          60
45      Glu Gln Leu Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile
      65          70          75
      Leu Lys Met Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu
      80          85          90          95
50      Phe Val Gln Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala
      100          105          110
      Ser Val Glu Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser
      115          120          125

```

EP 1 266 015 B1

Ala Thr Ser Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu  
130 135 140

Thr Val Ser  
145

<210> 43  
<211> 606  
<212> DNA  
<213> Artificial Sequence  
<220>  
<221> misc\_feature  
<223> Novel Sequence  
<400> 43

atgaaagcgt taacggccag gcaacaagag gtgtttgac tcattccgtga tcacatcagc 60  
cagacaggta tgccgccgac gcgtgcggaa atcgccgagc gtttgggggtt ccgttcccca 120  
aacgcggctg aagaacatct gaaggcgctg gcacgcaaaag gcgttattga aattgtttcc 180  
ggcgcatcac gcgggattcg tctgttcgag gaagagggaag aagggttgcc gctggtaggt 240  
cgtgtggctg ccggtgaacc acttctggcg caacagcata ttgaaggtca ttatcaggtc 300  
gatccttctt tattcaagcc gaatgctgat ttctgtctgc gcgtcagcgg gatgtcgatg 360  
aaagatatcg gcattatgga tgggtgacttg ctggcagtg ataaaaacta ggatgtacgt 420  
aacggtcagg tcgttgcgc acgtattgat gacgaagtta ccgttaagcg cctgaaaaaa 480  
cagggaata aagtcgaact gttgccagaa aatagcgagt ttaaaccaat tgtcgtagat 540  
cttcgtcagc agagcttcac cattgaaggg ctggcggttg gggttattcg caacggcgac 600  
tggtctg 606

<210> 44,  
<211> 202  
<212> PRT  
<213> Artificial Sequence  
<220>  
<221> misc feature  
<223> Novel Sequence  
<400> 44

Met Lys Ala Leu Thr Ala Arg Gln Gln Glu Val Phe Asp Leu Ile Arg  
1 5 10 15  
Asp His Ile Ser Gln Thr Gly Met Pro Pro Thr Arg Ala Glu Ile Ala  
20 25 30  
Gln Arg Leu Gly Phe Arg Ser Pro Asn Ala Ala Glu Glu His Leu Lys  
35 40 45  
Ala Leu Ala Arg Lys Gly Val Ile Glu Ile Val Ser Gly Ala Ser Arg  
50 55 60

EP 1 266 015 B1

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55

Gly Ile Arg Leu Leu Gln Glu Glu Glu Glu Gly Leu Pro Leu Val Gly  
65 70 75 80  
Arg Val Ala Ala Gly Glu Pro Leu Leu Ala Gln Gln His Ile Glu Gly  
85 90 95  
His Tyr Gln Val Asp Pro Ser Leu Phe Lys Pro Asn Ala Asp Phe Leu  
100 105 110  
Leu Arg Val Ser Gly Met Ser Met Lys Asp Ile Gly Ile Met Asp Gly  
115 120 125  
Asp Leu Leu Ala Val His Lys Thr Gln Asp Val Arg Asn Gly Gln Val  
130 135 140  
Val Val Ala Arg Ile Asp Asp Glu Val Thr Val Lys Arg Leu Lys Lys  
145 150 155 160  
Gln Gly Asn Lys Val Glu Leu Leu Pro Glu Asn Ser Glu Phe Lys Pro  
165 170 175  
Ile Val Val Asp Leu Arg Gln Gln Ser Phe Thr Ile Glu Gly Leu Ala  
180 185 190  
Val Gly Val Ile Arg Asn Gly Asp Trp Leu  
195 200

<210> 45  
<211> 271  
<212> DNA  
<213> Artificial Sequence  
<220>  
<221> misc\_feature  
<223> Novel Sequence  
<400> 45

atggggcccta aaaagaagcg taaagtcgcc cccccgaccg atgtcagcct gggggacgag 60  
ctccaccttag acggcgagga cgtggcgatg gcgcgatgcc acgcgctaga cgatttcgat 120  
ctggacatgt tgggggacgg ggattccccg gggccgggat ttacccccca cgactccgcc 180  
ccctacggcg ctctggatat ggccgacttc gagtttgagc agatgtttac cgatgccctt 240  
ggaattgacg agtacgggtg ggaattcccc g 271

<210> 46  
<211> 90  
<212> PRT  
<213> Artificial Sequence  
<220>  
<221> misc feature  
<223> Novel Sequence  
<400> 46

Met Gly Pro Lys Lys Lys Arg Lys Val Ala Pro Pro Thr Asp Val Ser  
1 5 10 15  
Leu Gly Asp Glu Leu His Leu Asp Gly Glu Asp Val Ala Met Ala His

EP 1 266 015 B1

		20		25		30	
	Ala Asp	Ala	Leu Asp Asp	Phe Asp	Leu Asp	Met	Leu Gly Asp Gly Asp
		35		40			45
5	Ser	Pro Gly	Pro Gly Phe	Thr	Pro His	Asp Ser	Ala Pro Tyr Gly Ala
		50		55			60
	Leu Asp	Met	Ala Asp	Phe	Glu Phe	Glu Gln	Met Phe Thr Asp Ala Leu
		65		70		75	80
10	Gly Ile	Asp Glu	Tyr Gly	Gly Glu	Phe	Pro	
			85		90		

<210> 47  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <221> misc feature  
 <223> Novel Sequence  
 <400> 47  
 ggagtactgt cctccgagc 19  
 <210> 48  
 <211> 666  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <221> misc feature  
 <223> Novel Sequence  
 <400> 48

	ggatcccccag	cttggaaattc	gacagggttat	cagcaacaac	acagtcatat	ccattctcaa	60
35	ttagctctac	cacagtgtgt	gaaccaatgt	atccagcacc	acctgtaacc	aaaacaattt	120
	tagaagtact	ttcactttgt	aactgagctg	tcattttatat	tgaattttca	aaaattctta	180
	cttttttttt	ggatggagcg	aaagaagttt	aataatcata	ttacatggca	ttaccaccat	240
40	atacatatcc	atatacatat	ccatatctaa	tcttaacctcg	actgctgtat	ataaaaccag	300
	tggttatatg	tacagtactg	ctgtatatata	aaccagtggt	tatatgtaca	gtacgtcgac	360
	tgctgtatat	aaaaccagtg	gttatatgta	cagtactgct	gtatataaaa	ccagtggtta	420
45	tatgtacagt	acgtcgaggg	atgataatgc	gattagtttt	ttagccttat	tctcggggtta	480
	attaatcagc	gaagcgatga	tttttgatct	attaacagat	atataaatgc	aaaaactgca	540
	taaccacttt	aactaatact	ttcaacattt	tccggtttgta	ttactttctta	ttcaaatgta	600
50	ataaaagtat	caacaaaaaa	ttgttaatat	acctctatac	tttaacgtca	aggagaaaaa	660
	actata						666

<210> 49  
 <211> 1542  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <221> misc feature

&lt;223&gt; Novel Sequence

&lt;400&gt; 49

5	ctggacctga aacacgaagt ggcttaccga ggggtgctec caggccaggt gaaggccgaa	60
	ccgggggtcc acaacggcca ggtcaacggc cactgagagg actgcatggc aggcggcgct	120
	ggtgccaat cgccgtctcc gggagcgggt gctcaacccc agcctaacaa tgggtattcg	180
10	tcgccactct cctcgggaag ctacgggccc tacagtccaa atgggaaat aggcctgtag	240
	gaactgtcgc cagcttcaag tataaatggg tgcagtacag atggcgaggc acgacgtcag	300
	aagaagggcc ctgcgccccg tcagcaagag gaactgtgtc tggtatcgcg ggacagagcc	360
	tccggatacc actacaatgc gctcacgtgt gaaggggtga aagggttctt cagacggagt	420
15	gttaccaaaa atcggttcta tatttgtaaa ttcggtcacg cttgcgaaat ggacatgtac	480
	atgcgacgga aatgccagga gtgccgctgt aagaagtgtc tagctgtagg catgaggcct	540
	gagtgcttag tacccgagac tcagtgcgcc atgaagcgga aagagaagaa agcacagaa	600
20	gagaaggaca aactgcctgt cagcacgacg acggtggacg accacatgcc gccattatg	660
	cagtgtgaac ctccacctcc tgaagcagca aggattcacg aagtggtecc aaggtttctc	720
	tccgacaagc tgttggagac aaaccggcag aaaaacatcc ccagttgac agccaaccag	780
25	cagttcctta tcgccaggct catctggtac caggacgggt acgagcagcc ttctgatgaa	840
	gatttgaaga ggattacgca gacgtggcag caagcggacg atgaaaacga agagtctgac	900
	actcccttcc gccagatcac agagatgact atcctcacgg tccaacttat cgtggagttc	960
30	gcgaagggat tgcagggtt cgccaagatc tcgcagcctg atcaaatatc gctgcttaag	1020
	gcttgctcaa gtgaggtaat gatgctccga gtcgcgcgac gatacatgac ggccctcagac	1080
	agtgttctgt tcgcgaacaa ccaagcgtac actcgcgaca actaccgcaa ggctggcatg	1140
35	gcctacgtca tcgaggatct actgcacttc tgcgggtgca tgtactctat ggcgttggac	1200
	aacatccatt acgcctgtct caccgctgtc gtcacttttt ctgaccggcc aggggtggag	1260
	cagccgcaac tgggtggaaga aatccagcgg tactacctga atacgctccg catctatata	1320
40	ctgaaccacg tgagcgggtc ggcgcgttcg tccgtcatat acggcaagat cctctcaatc	1380
	ctctctgagc tacgcacgct cggcatgcaa aactccaaca tgtgcatctc cctcaagctc	1440
	aagaacagaa agctcgccgc ttctcctcag gagatctggg atgtggcgga catgtcgcac	1500
45	acccaaccgc cgcctatcct cgagtccccc acgaatctct ag	1542

&lt;210&gt; 50

&lt;211&gt; 513

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; misc feature

&lt;223&gt; Novel Sequence

&lt;400&gt; 50

EP 1 266 015 B1

1 Leu Asp Leu Lys His Glu Val Ala Tyr Arg Gly Val Leu Pro Gly Gln  
 5 Val Lys Ala Glu Pro Gly Val His Asn Gly Gln Val Asn Gly His Val  
 10 Arg Asp Trp Met Ala Gly Gly Ala Gly Ala Asn Ser Pro Ser Pro Gly  
 15 Ala Val Ala Gln Pro Gln Pro Asn Asn Gly Tyr Ser Ser Pro Leu Ser  
 20 Ser Gly Ser Tyr Gly Pro Tyr Ser Pro Asn Gly Lys Ile Gly Arg Glu  
 25 Glu Leu Ser Pro Ala Ser Ser Ile Asn Gly Cys Ser Thr Asp Gly Glu  
 30 Ala Arg Arg Gln Lys Lys Gly Pro Ala Pro Arg Gln Gln Glu Glu Leu  
 35 Cys Leu Val Cys Gly Asp Arg Ala Ser Gly Tyr His Tyr Asn Ala Leu  
 40 Thr Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Ser Val Thr Lys Asn  
 45 Ala Val Tyr Ile Cys Lys Phe Gly His Ala Cys Glu Met Asp Met Tyr  
 50 Met Arg Arg Lys Cys Gln Glu Cys Arg Leu Lys Lys Cys Leu Ala Val  
 55 Gly Met Arg Pro Glu Cys Val Val Pro Glu Thr Gln Cys Ala Met Lys  
 60 Arg Lys Glu Lys Lys Ala Gln Lys Glu Lys Asp Lys Leu Pro Val Ser  
 65 Thr Thr Thr Val Asp Asp His Met Pro Pro Ile Met Gln Cys Glu Pro  
 70 Pro Pro Pro Glu Ala Ala Arg Ile His Glu Val Val Pro Arg Phe Leu  
 75 Ser Asp Lys Leu Leu Glu Thr Asn Arg Gln Lys Asn Ile Pro Gln Leu  
 80 Thr Ala Asn Gln Gln Phe Leu Ile Ala Arg Leu Ile Trp Tyr Gln Asp  
 85 Gly Tyr Glu Gln Pro Ser Asp Glu Asp Leu Lys Arg Ile Thr Gln Thr  
 90 Trp Gln Gln Ala Asp Asp Glu Asn Glu Glu Ser Asp Thr Pro Phe Arg

EP 1 266 015 B1

	290		295		300	
	Gln Ile Thr Glu Met Thr	Ile Leu Thr Val	Gln Leu Ile Val	Glu Phe		
	305	310	315	320		
5	Ala Lys Gly Leu Pro Gly Phe Ala Lys	Ile Ser Gln Pro Asp	Gln Ile			
	325	330	335			
	Thr Leu Leu Lys Ala Cys Ser Ser Glu Val Met Met Leu Arg Val Ala					
	340	345	350			
10	Arg Arg Tyr Asp Ala Ala Ser Asp Ser Val Leu Phe	Ala Asn Asn Gln				
	355	360	365			
	Ala Tyr Thr Arg Asp Asn Tyr Arg Lys Ala Gly Met	Ala Tyr Val Ile				
	370	375	380			
15	Glu Asp Leu Leu His Phe Cys Arg Cys Met Tyr Ser Met Ala Leu Asp					
	385	390	395			
	Asn Ile His Tyr	Ala Leu Leu Thr Ala Val Ile Phe Ser Asp Arg				
	405	410	415			
20	Pro Gly Leu Glu Gln Pro Gln Leu Val Glu Glu Ile Gln Arg Tyr Tyr					
	420	425	430			
	Leu Asn Thr Leu Arg Ile Tyr	Ile Leu Asn Gln Leu Ser Gly Ser Ala				
	435	440	445			
25	Arg Ser Ser Val Ile Tyr Gly Lys Ile Leu Ser	Ile Leu Ser Glu Leu				
	450	455	460			
	Arg Thr Leu Gly Met Gln Asn Ser Asn Met Cys Ile Ser Leu Lys Leu					
	465	470	475			
30	Lys Asn Arg Lys Leu Pro Pro Phe Leu Glu Glu Ile Trp Asp Val Ala					
	485	490	495			
	Asp Met Ser His Thr Gln Pro Pro Pro Ile Leu Glu Ser Pro Thr Asn					
	500	505	510			
35	Leu					
	<210> 51					
	<211> 4375					
	<212> DNA					
40	<213> Artificial Sequence					
	<220>					
	<221> misc_feature					
	<223> Novel Sequence					
	<400> 51					
45						
	tgtaattttg atggggcgccg tgcgcaccg tgtgccatat tgccatccag tcgaatagaa				60	
	aaaaaaaaa aaaaaaaaa atcagttgtt ttgtccctcg ctcgctttcg agtgtattcg				120	
50	gaatattaga cgtcataatt cacgagtgc ttttaaatat atatagcgat tagcggggcc				180	
	gtttgttggg cgtgcgcttg cgtttagtgg agtgcagggg tagtgaggcg agtatggtag				240	
	ttcgtggtca tgtcaagtgt ggcgaagaaa gacaagccga cgatgtcggt gacggcgctg				300	
55						

	atcaactggg	cgccggccggc	gccgccagc	ccgccgcagc	cgagtcagc	gtcgctcg	360
	ccggcagcca	tgctgcagca	gctccgacg	cagtcgaatgc	agtcgtttaa	ccacatccca	420
5	actgtcgatt	gctcgctcga	tatgcagtgg	cttaatttag	aacctgagtt	catgtcgct	480
	atgtcaacct	ctgagatgaa	accagacacc	gccatgcttg	atggggtacg	agacgacgcc	540
	acttcgccc	ctaacttcaa	gaactacccg	cctaatacacc	ccctgagtgg	ctccaaacac	600
10	ctatgtctta	tatgcggcga	cagggcgctc	gggaagcact	atggggtgta	cagtgtcgaa	660
	ggatgcaagg	gtttcttcaa	gcggaccgtc	cggaaggacc	tgctgtacgc	ttgccgggag	720
	gagcggaact	gcatacataga	caagcgacaa	aggaaccgat	gccagtagtc	ccgctatcaa	780
15	aagtgtttgg	cttgcggtat	gaagcgagag	gcggtgcaag	aggagcgcca	gaggaatgct	840
	cgccggccgg	aggatgcgca	cccagtagc	tcggtgcagg	taagcgatga	gctgtcaatc	900
	gagcgccata	cgagatgga	gtctttggtg	gcagatccca	gcgaggagtt	ccagttcctc	960
20	cgcgtagggc	ctgacagcaa	cgtgcctcca	cgttaccgcg	cgcccgcttc	ctccctctgc	1020
	caaataggca	acaagcaaat	agcggcggtg	gtggtatggg	cgccgcacat	ccctcatttc	1080
	gggcagctgg	agctggacga	tcaagtggta	ctcatcaagg	cctcctggaa	tgagctgcta	1140
25	ctcttcgcc	tcgcttgccg	ctctatggag	tatttggaag	atgagagggg	gaacggggac	1200
	ggaacgcgga	gcaccactca	gccacaactg	atgtgtctca	tgcttgccat	gacgttgcat	1260
	cgcaactcgg	cgcagcaggc	ggcggtgggc	gccatcttcg	accgcgtgct	gtccgagctc	1320
	agtcggaaga	tgcgacactt	gcgcatggac	caggccgagt	acgtcgcgct	caaagccatc	1380
30	gtgtgctca	accctgatgt	gaaaggactg	aagaatcgcc	aagaagttag	cgttttgca	1440
	gaaaaaatgt	tctcttgctc	ggacgactac	tgccggcggt	cgcgaaagca	cgaggaaggc	1500
	cggtttgcgt	ccttctgctc	gcggctgcca	gctctccgct	ccatctcgct	caagagcttc	1560
35	gaacacctct	acttcttcca	cctcgtggcc	gaaggctcca	tcagcggata	catagcagag	1620
	gcgctccgaa	accacgcgcc	tcgcatcgac	gtcaatgcca	tgatgtaaa	tgcatataac	1680
	gccctgccga	tgtgagaaga	actatggcta	atagaagcga	aactgaatac	atctaggggtg	1740
40	ggacttaact	tgggactatc	attaaagtat	cacgcaaatt	atgcgtagtc	agaaagtgc	1800
	gtcgatcaaa	cttttttata	aacgaattga	gtttctaacg	actgcaacac	agcggagttt	1860
	tgcttctgat	agtttttatt	ctaattggta	agatgcttta	cacgggcatt	attgacattc	1920
45	aagtgtgaag	ggaagtgtac	aaccttgaca	tttatatcac	gtttgtaatt	ggttaataaa	1980
	attaatat	cacaagtaa	actaacatca	acgtcacgat	actaacgcca	tttagtgata	2040
	ttttctatg	caagaaactc	attgttttga	taaaaatttt	ttctaattac	tcagtgaaac	2100
50	tcataccaa	gtgaccagtc	ttcccgacga	gttgcccggtg	taaaatcctc	tttagggaca	2160
	tatcccccg	tatctcatga	aattccaagg	atcagtaggg	gccaatcccc	ccagtggtt	2220
	gggaggcaga	attttcgata	atctacgact	attgttagcc	tacgaattag	ttgaattttt	2280

55



	tgaaattatt	tttattaagt	cgccactttc	caaacacatc	agcagggtat	atgtgcaatt	2340
	ttgtaacgat	aaactctatc	atttctgata	tttatcgaaa	ttttatctta	cataacatgc	2400
5	tggctggctc	agggtgtttg	tagttacata	tgtatctacg	gtttgtttta	aattatagct	2460
	ttttttattg	aatctgtata	aaattgagtt	atcttacttc	acactacgat	cgagtaaacc	2520
	catcgtcagc	tacgaaaaac	taatcgtata	aggcgttaaga	gtaaaatact	aattgacaac	2580
10	cagcaacgag	gaccacctca	gtcctcgtgc	ttacattgtg	cgtagctta	atatgatgga	2640
	agctgtcgtc	gttacgacat	tagataaagt	gcatgaatac	caaaaaatga	ccatcccgtg	2700
	ctgatctctc	atgctctcgc	tgcgtgggac	ccgtgtcgag	tgtcgttaag	actgactaat	2760
15	attttagact	aggcgtctat	gcttcagtaa	ttccttatac	atattataag	tcattccaaat	2820
	aacgagtaag	gcgcatgtt	gagatcagca	ttccgagagt	caaagagccc	ctaactgtac	2880
	tgagaagtag	agacaatata	ctgattttct	gagatgaacg	caaccgagat	tgacactaaa	2940
20	aatctattta	tggatttcaa	aatggcgatg	cttgattgtc	tgcggcgtgg	atagactgaa	3000
	atgggttttg	ttaacactgg	atattgtttt	tattagttaa	tagtcttaca	ttgcaagtgtg	3060
	gtaattcggg	gctaataatc	accggtttgt	taactatcta	acggttccca	gtgtcaggca	3120
25	cacatcttct	ccaagcagac	aacgcaagag	tgtacaaaa	gtacatgtta	caaaaataag	3180
	aacattcgtc	ggataagtgt	aacagttgat	aggtaaaaga	aatggggccg	cctctttatt	3240
	attacgtagc	cgtaaaaata	ttaacgtatt	tagtttagat	gttcagctaa	ttaggataat	3300
30	tctatttgtc	gagtaacatg	atgtccatag	tgaattaata	taataattag	actgttacgc	3360
	gtaggtaatt	ataaagttta	ccaaatctct	cttcaaagca	aaaactttgt	acacttccgt	3420
	actgagacgt	cgtagcttat	tctgattcac	gaaatatttg	gatcacattg	ttacaaggcg	3480
35	accgtcacgt	agtatatgat	tatttacaaa	tgacacgtat	gtatcaatgc	tataagtgtt	3540
	ttcgtttacat	atgtcgtgtc	tttaacgtgc	atttcgatgt	gcagattaaa	aatagcaaga	3600
	aatcttgaaa	ttgttttaga	aaatatttga	tttccttatt	gaaagtatt	tttaaatgta	3660
	aatatttctg	aatcataata	attatgtatt	gtgtagttat	ttcaccttta	cggttgggat	3720
40	attattttaa	ggtggcctac	gaaagtgtat	ataaccatcc	gcgtcctcaa	aaaggccagt	3780
	ttatttttgt	acctcatata	tactaattac	gtaagtaata	tcaggcgaa	ggttgactaa	3840
	caactaacca	gtattaaaaa	ttaaaagact	tcgtccta	aaaatgtaat	atctatgtat	3900
45	aaaaatgaaa	aatctggcgt	ataataggtg	aaattaaact	agattgttaa	tgaatgtgat	3960
	gtctcataaa	cgtttagttt	ttaatgagaa	acatgtttag	tcgcctacta	taagacgaga	4020
	cggcaagctc	accgagttaa	ctcgtaaaca	ggaatgttga	aaaagatgac	acaatttata	4080
50	tttggtattg	aaattatgac	taaccatgag	ctctatcggt	tgttatggat	gcatagtatt	4140
	gctgttgaaa	ataatggaat	taggtaatta	ctgcattaat	gttgaaaact	tgatatattt	4200

55

```

ctatggttgg gtatgaattc tatgttgaa gtgttcagc ggttgtaaag atgatttata 4260
atgatgttca ctaaataatct gactaaatgt aagtatttt tttttgtata gacatagctt 4320
taagatgaag gtgattaaac tttatcctta tcacaataaa aaaaaaaaaa aaaaa 4375

```

```

<210> 52
<211> 472
<212> PRT
<213> Artificial Sequence
<220>
<221> misc_feature
<223> Novel Sequence
<400> 52

```

```

Met Ser Ser Val Ala Lys Lys Asp Lys Pro Thr Met Ser Val Thr Ala
1      5      10      15
Leu Ile Asn Trp Ala Arg Pro Ala Pro Pro Gly Pro Pro Gln Pro Gln
20     20     25     30
Ser Ala Ser Pro Ala Pro Ala Ala Met Leu Gln Gln Leu Pro Thr Gln
35     40     45
Ser Met Gln Ser Leu Asn His Ile Pro Thr Val Asp Cys Ser Leu Asp
50     55     60
Met Gln Trp Leu Asn Leu Glu Pro Gly Phe Met Ser Pro Met Ser Pro
65     70     75     80
Pro Glu Met Lys Pro Asp Thr Ala Met Leu Asp Gly Leu Arg Asp Asp
85     90     95
Ala Thr Ser Pro Pro Asn Phe Lys Asn Tyr Pro Pro Asn His Pro Leu
100    105    110
Ser Gly Ser Lys His Leu Cys Ser Ile Cys Gly Asp Arg Ala Ser Gly
115    120    125
Lys His Tyr Gly Val Tyr Ser Cys Glu Gly Cys Lys Gly Phe Phe Lys
130    135    140
Arg Thr Val Arg Lys Asp Leu Ser Tyr Ala Cys Arg Glu Glu Arg Asn
145    150    155    160
Cys Ile Ile Asp Lys Arg Gln Arg Asn Arg Cys Gln Tyr Cys Arg Tyr
165    170    175
Gln Lys Cys Leu Ala Cys Gly Met Lys Arg Glu Ala Val Gln Glu Glu
180    185    190
Arg Gln Arg Asn Ala Arg Gly Ala Glu Asp Ala His Pro Ser Ser Ser
195    200    205
Val Gln Val Ser Asp Glu Leu Ser Ile Glu Arg Leu Thr Glu Met Glu
210    215    220
Ser Leu Val Ala Asp Pro Ser Glu Glu Phe Gln Phe Leu Arg Val Gly
225    230    235    240
Pro Asp Ser Asn Val Pro Pro Arg Tyr Arg Ala Pro Val Ser Ser Leu

```

EP 1 266 015 B1

245 250 255  
 Cys Gln Ile Gly Asn Lys Gln Ile Ala Ala Leu Val Val Trp Ala Arg  
 260 265 270  
 5 Asp Ile Pro His Phe Gly Gln Leu Glu Leu Asp Asp Gln Val Val Leu  
 275 280 285  
 Ile Lys Ala Ser Trp Asn Glu Leu Leu Leu Phe Ala Ile Ala Trp Arg  
 290 295 300  
 10 Ser Met Glu Tyr Leu Glu Asp Glu Arg Glu Asn Gly Asp Gly Thr Arg  
 305 310 315 320  
 Ser Thr Thr Gln Pro Gln Leu Met Cys Leu Met Pro Gly Met Thr Leu  
 325 330 335  
 15 His Arg Asn Ser Ala Gln Gln Ala Gly Val Gly Ala Ile Phe Asp Arg  
 340 345 350  
 Val Leu Ser Glu Leu Ser Leu Lys Met Arg Thr Leu Arg Met Asp Gln  
 355 360 365  
 20 Ala Glu Tyr Val Ala Leu Lys Ala Ile Val Leu Leu Asn Pro Asp Val  
 370 375 380  
 Lys Gly Leu Lys Asn Arg Gln Glu Val Asp Val Leu Arg Glu Lys Met  
 385 390 395 400  
 25 Phe Ser Cys Leu Asp Asp Tyr Cys Arg Arg Ser Arg Ser Asn Glu Glu  
 405 410 415  
 Gly Arg Phe Ala Ser Leu Leu Leu Arg Leu Pro Ala Leu Arg Ser Ile  
 420 425 430  
 30 Ser Leu Lys Ser Phe Glu His Leu Tyr Phe Phe His Leu Val Ala Glu  
 435 440 445  
 Gly Ser Ile Ser Gly Tyr Ile Arg Glu Ala Leu Arg Asn His Ala Pro  
 450 455 460  
 35 Pro Ile Asp Val Asn Ala Met Met  
 465 470

<210> 53  
 <211> 1404  
 40 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <221> misc feature  
 <223> Novel Sequence  
 45 <400> 53

atggacacca aacatttctt gccgtctgac ttctctaccc aggtgaactc ttcgtccctc 60  
 aactctccaa cgggtcgagg ctccatggct gtcccctcgc tgcacccctc ctggggtccg 120  
 50 ggaatcggct ctccactggg ctgcctggg cagctgcact ctctatcag caccctgagc 180  
 tccccatca atggcatggg tccgcccttc tctgtcatca gctcccccat gggcccgcac 240  
 tccatgtcgg taccaccac accacattg ggttcggga ctggtagccc ccagctcaat 300  
 55

tcacctcatga accctgtgag cagcactgag gatatacaagc cgccactagg cctcaatggc 360  
 gtctctcaagg ttccctgcccc tccctcagga aatatggcct ccttcaccaa gcacatctgt 420  
 5 gctatctgtg gggaccgctc ctacaggcaaa cactatgggg tatacagttg tgagggctgc 480  
 aagggtctct tcaagaggac agtacgcaaa gacctgacct acacctgccg agacaacaag 540  
 gactgcctga tcgacaagag acagcggaac cgggtgcagt actgccgcta ccagaagtgc 600  
 10 ctggccatgg gcatgaagcg ggaagctgtg caggaggagc ggcagcgggg caaggaccgg 660  
 aatgagaacg aggtggagtc caccagcagt gccaacgagg acatgcctgt agagaagatt 720  
 ctggaagccg agcttctgtg cgagcccaag actgagacat acgtggaggc aaacatgggg 780  
 15 ctgaacccca gctcaccaaa tgaccctgtt accaaccatct gtcaagcagc agacaagcag 840  
 ctcttctactc ttgtggagtg ggccaagagg atcccacact tttctgagct gccctagac 900  
 gaccaggtea tcctgctacg ggcaggctgg aacgagctgc tgatcgccctc ctctcccac 960  
 20 cgctccatag ctgtgaagaa tgggattctc ctggccaccg gcctgcacgt acaccggaac 1020  
 agcgctcaca gtgtgggggt gggcgccatc ttgacaggg tgctaacaga gctggtgtct 1080  
 aagatgcgtg acatgcagat ggacaagacg gagctggggt gcctgcgagc cattgtcctg 1140  
 25 ttcaaccctg actctaaggg gctctcaaac cctgctgagg tggaggcggt gagggagaag 1200  
 gtgtatcgct cactagaagc gtactgcaaa cacaagtacc ctgagcagcc gggcagggtt 1260  
 gccaaagctgc tgctccgctt gcctgacctg cgttccatcg ggctcaagtg cctggagcac 1320  
 ctgttcttct tcaagctcat cggggacacg cccatcgaca ccttcctcat ggagatgctg 1380  
 30 gaggcaccac atcaagccac ctac 1404

&lt;210&gt; 54

&lt;211&gt; 467

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; misc feature

&lt;223&gt; Novel Sequence

&lt;400&gt; 54

Met Asp Thr Lys His Phe Leu Pro Leu Asp Phe Ser Thr Gln Val Asn  
 1 5 10 15  
 45 Ser Ser Ser Leu Asn Ser Pro Thr Gly Arg Gly Ser Met Ala Val Pro  
 20 25 30  
 Ser Leu His Pro Ser Leu Gly Pro Gly Ile Gly Ser Pro Leu Gly Ser  
 35 40 45  
 50 Pro Gly Gln Leu His Ser Pro Ile Ser Thr Leu Ser Ser Pro Ile Asn  
 55 60  
 Gly Met Gly Pro Pro Phe Ser Val Ile Ser Ser Pro Met Gly Pro His  
 65 70 75 80

Ser Met Ser Val Pro Thr Thr Pro Thr Leu Gly Phe Gly Thr Gly Ser  
 85 90 95  
 5 Pro Gln Leu Asn Ser Pro Met Asn Pro Val Ser Ser Thr Glu Asp Ile  
 100 105 110  
 Lys Pro Pro Leu Gly Leu Asn Gly Val Leu Lys Val Pro Ala His Pro  
 115 120 125  
 10 Ser Gly Asn Met Ala Ser Phe Thr Lys His Ile Cys Ala Ile Cys Gly  
 130 135 140  
 Asp Arg Ser Ser Gly Lys His Tyr Gly Val Tyr Ser Cys Glu Gly Cys  
 145 150 155 160  
 15 Lys Gly Phe Phe Lys Arg Thr Val Arg Lys Asp Leu Thr Tyr Thr Cys  
 165 170 175  
 Arg Asp Asn Lys Asp Cys Leu Ile Asp Lys Arg Gln Arg Asn Arg Cys  
 180 185 190  
 20 Gln Tyr Cys Arg Tyr Gln Lys Cys Leu Ala Met Gly Met Lys Arg Glu  
 195 200 205  
 Ala Val Gln Glu Glu Arg Gln Arg Gly Lys Asp Arg Asn Glu Asn Glu  
 210 215 220  
 25 Val Glu Ser Thr Ser Ser Ala Asn Glu Asp Met Pro Val Glu Lys Ile  
 225 230 235 240  
 Leu Glu Ala Glu Leu Ala Val Glu Pro Lys Thr Glu Thr Tyr Val Glu  
 245 250 255  
 30 Ala Asn Met Gly Leu Asn Pro Ser Ser Pro Asn Asp Pro Val Thr Asn  
 260 265 270  
 Ile Cys Gln Ala Ala Asp Lys Gln Leu Phe Thr Leu Val Glu Trp Ala  
 275 280 285  
 35 Lys Arg Ile Pro His Phe Ser Glu Leu Pro Leu Asp Gln Val Ile  
 290 295 300  
 Leu Leu Arg Ala Gly Trp Asn Glu Leu Leu Ile Ala Ser Phe Ser His  
 305 310 315 320  
 40 Arg Ser Ile Ala Val Lys Asp Gly Ile Leu Leu Ala Thr Gly Leu His  
 325 330 335  
 Val His Arg Asn Ser Ala His Ser Ala Gly Val Gly Ala Ile Phe Asp  
 340 345 350  
 45 Arg Val Leu Thr Glu Leu Val Ser Lys Met Arg Asp Met Gln Met Asp  
 355 360 365  
 Lys Thr Glu Leu Gly Cys Leu Arg Ala Ile Val Leu Phe Asn Pro Asp  
 370 375 380  
 Ser Lys Gly Leu Ser Asn Pro Ala Glu Val Glu Ala Leu Arg Glu Lys  
 385 390 395 400  
 50 Val Tyr Ala Ser Leu Glu Ala Tyr Cys Lys His Lys Tyr Pro Glu Gln  
 405 410 415  
 Pro Gly Arg Phe Ala Lys Leu Leu Leu Arg Leu Pro Ala Leu Arg Ser

# EP 1 266 015 B1

		420		425		430	
	Ile Gly	Leu Lys Cys <sup>1</sup> Leu Glu	His	Leu Phe Phe Phe	Lys	Leu Ile Gly	
5		435		440		445	
	Asp Thr	Pro Ile Asp Thr	Phe	Leu Met Glu Met	Leu	Glu Ala Pro His	
		450		455		460	
10	Gln Ala Thr						
		465					
	<210> 55						
	<211> 309						
	<212> DNA						
15	<213> Artificial Sequence						
	<220>						
	<221> misc_feature						
	<223> Novel Sequence						
	<400> 55						
20							
	ggtgtgga	gtccccaggc	tccccagcag	gcagaagtat	gcaaagcatg	catctcaatt	60
	agtcagaca	caggtgtgga	aagtccccagc	gctccccagc	aggcagaagt	atgcaaagca	120
25	tgcatctcaa	ttagtcagca	accatagtcc	cgccccatac	tccgcccatc	ccgccctaa	180
	ctccgccag	tccgcccatc	tctccgcccc	atggctgact	aatttttttt	atttatgcag	240
	aggccgaggc	cgctccggcc	tctgagctat	tccagaagta	gtgaggaggc	ttttttggag	300
30	gcctaggct						309
	<210> 56						
	<211> 24						
	<212> DNA						
35	<213> Artificial Sequence						
	<220>						
	<221> misc feature						
	<223> Novel-Sequence						
	<400> 56						
40	tatataatgg	atccccgggt	accg	24			
	<210> 57						
	<211> 1653						
	<212> DNA						
	<213> Artificial Sequence						
45	<220>						
	<221> misc feature						
	<223> Novel Sequence						
	<400> 57						
50	atggaagacg	ccaaaaacat	aaagaaaggc	cggcgcccat	tctatcctct	agaggatgga	60
	accgctggag	agcaactgca	taaggctatg	aagagatacg	cctggtttcc	tggacaatt	120
55	gcttttacag	atgcacatat	cgaggatgaac	atcacgtacg	cgaataactt	cgaatatgtcc	180

	gttcggttgg cagaagctat gaaacgatat gggctgaata caaatcacag aatcgtcgta	240
5	tgcaagtgaat actctcttca attctttatg ccggtgttgg gcgcgttatt tatcggagtt	300
	gcagttgcgc ccgcgaacga catttataat gaacgtgaat tgctcaacag tatgaacatt	360
	tcgcagccta ccgtagtgtt tgtttccaaa aaggggttgc aaaaaatttt gaacgtgcaa	420
10	aaaaaattac caataatcca gaaaattatt atcatggatt ctaaaacgga ttaccaggga	480
	tttcagtcga tgtacacggt cgtcacatct catctacctc ccggttttaa tgaatacgat	540
	tttgtaccag agtcctttga tcgtgacaaa acaattgcac tgataatgaa ttccctctgga	600
15	tctactgggt tacctaagggt tgtggccctt ccgcatagaa ctgcctgcgt cagattctcg	660
	catgccagag atcctatttt tggcaatcaa atcattccggt atactcgat tttaaagtgtt	720
	gttccattcc atcacggttt tggaaatggtt actacactcg gatatttgat atgtggattt	780
20	cgagtcgtct taatgtatag atttgaagaa gagctgtttt tacgatccct tcaggattac	840
	aaaaattcaa gtgcgttgct agtaccaacc ctattttcat tcttcgcaa aagcactctg	900
	attgacaaat acgattttat taatttacac gaaattgctt ctgggggcgc acctcttctg	960
25	aaagaagtcg gggaagcgggt tgcaaaacgc ttccatcttc cagggatagc acaaggatat	1020
	gggctcactg agactacatc agctattctg attacaccgc agggggatga taaaccgggc	1080
	gcggtcggtt aagtgtgtcc attttttgaa gcgaaggttg tggatctgga taccgggaaa	1140
30	acgctgggag ttaatcacag aggcgaatta tgtgtcagag gacctatgat tatgtccggt	1200
	tatgtaaaca atccggaagc gaccaacgcc ttgattgaca aggatggatg gctacattct	1260
	ggagacatag ctactctgga cgaagacgaa cacttcttca tagttgaccg ctggaagtct	1320
35	ttaattaaat acaaaggata tcagggtggcc ccgctgaat tggaatcgat attgttaca	1380
	cacccaaca tcttcgacgc gggcgtggca ggtcttcccg acgatgacgc cggatgaactt	1440
40	ccgcgcgcgc ttgttgtttt ggagcacgga aagacgatga cggaaaaaga gatcgtggat	1500
	tacgtcgcca gtcaagtaac aaccgcgaaa aagttgcgcg gaggagttgt gtttgtggac	1560
	gaagtaccga aaggtcttac cggaaaactc gacgcaagaa aaatcagaga gatcctcata	1620
	aaggccaaga agggcggaat gtccaaattg taa	1653

45	<210> 58
	<211> 867
	<212> DNA
	<213> Artificial Sequence
	<220>
50	<221> misc feature
	<223> Novel Sequence
	<400> 58

aagcgagagg cgggtgcaaga ggagcgccag aggaatgctc gcgcgcgga ggatgcgcac 60

55

EP 1 266 015 B1

	ccgagtagct cgggtgcaggt aagcgatgag ctgtcaatcg agcgcctaac ggagatggag	120
	tctttggtgg cagatccca gaggaggttc cagttcctco gcgtggggcc tgacagcaac	180
5	gtgcctccac gttaccgcgc gcccgctctcc tccctctgcc aaataggcaa caagcaata	240
	cgggcggttg tggtagggc gcgcgacatc cctcatttcg ggcaagctga gctggacgat	300
	caagtggtag tcatcaagc ctccctggaat gagctgctac tcttcgccat cgcctggcgc	360
10	tctatggagt atttggaaga tgagagggag aacggggacg gaacgcggag caccactcag	420
	ccacaactga tgtgtctcat gcctggcatg acgttgccac gcaactcggc gcagcaggcg	480
	ggcgtagggc ccatcttcga ccgcgtgctg tccgagctca gtctgaagat gcgcaccttg	540
15	cgcatggacc aggccgagta cgtcgcgctc aaagccatcg tgctgctcaa cctgatgtg	600
	aaaggactga agaatacgga agaagttgac gttttgcgag aaaaaatgtt ctcttgcttg	660
	gacgactact gccggcggtc gcgaagcaac gaggaaggcc ggtttgcgtc ctgctgctg	720
20	cggtcgccag ctctccgctc catctcgctc aagagdttcg aacacctcta ctcttccac	780
	ctcgtggccg aaggctccat cagcggatag atacgagagc cgctccgaaa ccacgcgcct	840
	ccgatcgacg tcaatgccat gatgtaa	867

25  
 <210> 59  
 <211> 225  
 <212> DNA  
 <213> Artificial Sequence  
 30  
 <220>  
 <221> misc\_feature  
 <223> Novel Sequence  
 <400> 59

35	tcgacattgg acaagtgcac tgaacccttg tctctcgaga gacaaggggg ttcaatgcac	60
	ttgtccaatg tcgagagaca aggggggttca atgcacttgt ccaatgtcga gagacaagg	120
	ggttcaatgc acttgtccaa tgtcgagaga caaggggggt caatgcactt gtccaatgtc	180
40	gagagacaag ggggttcaat gcacttgtcc aatgtcgact ctaga	225

45  
 <210> 60  
 <211> 619  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <221> misc feature  
 <223> Novel Sequence  
 50  
 <400> 60

55	cgttacataa cttacggtaa atggcccgcg tggctgacgc cccaacgacc cccgccatt	60
	gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca	120
	atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc	180



EP 1 266 015 B1

5 aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta 240  
 catgacctta tgggactttc ctacttgcca gtacatctac gtattagtca tcgctattac 300  
 catggtgatg cgggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 360  
 atttccaagt ctccacccca ttgacgtcaa tggggagtttg ttttggcacc aaaatcaacg 420  
 ggactttcca aaatgtcgtg acaactccgc cccattgacg caaatgggag gtaggcgtgt 480  
 10 acggtgggag gtctatatata gcagagctcg tttagtgaac cgtcagatcg cctggagacg 540  
 ccatccacgc tgttttgacc tccatagaag acaccgggac cgatccagcc tccgcgggccg 600  
 ggaacgggtgc attggaacg 619

15 <210> 61  
 <211> 262  
 <212> DNA  
 <213> Artificial Sequence  
 20 <220>  
 <221> misc\_feature  
 <223> Novel Sequence  
 <400> 61

25 atgtagtctt atgcaatact cttgtagtct tgcaacatgg taacgatgag ttagcaacat 60  
 gccttacaag gagagaaaaa gcaccgtgca tgccgatagg tggaagtaag gtggtacgat 120  
 cgtgccttat taggaaggca acagacgggt ctgacatgga ttggacgaac cactgaattc 180  
 30 cgcaattgag agataattgta ttttaagtgcc tagctcgata caataaacgc catttgacca 240  
 ttcaccacat tggagtgcac ct 262

35 <210> 62  
 <211> 1247  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <221> misc feature  
 40 <223> Novel Sequence  
 <400> 62

45 tctatttctt caggccgtga ggaactgtcg ccagcttcaa gtataaatgg gtgcagtaca 60  
 gatggcgagg cagcacgtca gaagaaggcg cctgcgcccc gtcagcaaga ggaactgtgt 120  
 ctggtatgag gggacagagc ctccggatac cactacaatg cgctcactgt tgaagggtgt 180  
 aaagggttct tcagacggag tgttaccaaa aatgcggttt atatttgtta attcggtcac 240  
 50 gcttgcgaaa tggacatgta catgcgacgg aaatgccagg agtgccgcct gaagaagtgc 300  
 ttagctgtag gcatgaggcc tgagtgcgta gtaccgaga ctcagtgcgc catgaagcgg 360  
 aaagagaaga aagcacagaa ggagaaggac aaactgcctg tcagcacgac gacggtggac 420

55

gaccacatgc cgcccattat gcagtgtgaa cctccacetc ctgaagcagc aaggattcac 480  
gaagtgggtcc caaggtttct ctccgacaag ctgttggaga caaacccggca gaaaaacatc 540  
5 cccagtttga cagccaaacca gcagtttcct atcgccaggc tcattctggta ccaggacggg 600  
tacgagcagc cttctgatga agatttgaag aggattacgc agacgtggca gcaagcggac 660  
gatgaaaacg aagagtctga cactcccttc cgccagatca cagagatgac tatctctcagc 720  
10 gtccaaacta tcgtggagtt cgcgaaaggga ttgccagggt tcgccaagat ctgcgagcct 780  
gatcaaatta cgctgcttaa ggcttgctca agtgaggtaa tgatgctccg agtcgcgcga 840  
cgatacagatg cggcctcaga cagtgttctg ttgcggaaca accaagcgta cactcgcgac 900  
15 aactaccgca aggctggcat ggcttacgtc atcgaggatc tactgcactt ctgccgggtgc 960  
atgtacteta tggcggttga caacatccat tacgcgctgc tcacggctgt cgtcatcttt 1020  
tctgaccggc cagggttggga gcagccgcaa ctggtggaag aaatccagcg gtactacctg 1080  
20 aatacgctcc gcattctatat cctgaaccag ctgagcgggt cggcgcgctc gtccgtcata 1140  
tacggcaaga tcctctcaat cctctctgag ctacgcacgc tcggcatgca aaatcccaac 1200  
atgtgcatct cctcaagct caagaacaga aagctgccgc ctttcct 1247

<210> 63  
<211> 440  
<212> PRT  
<213> Artificial Sequence  
30 <220>  
<221> misc feature  
<223> Novel Sequence  
<400> 63

35 Ser Ile Ser Ser Gly Arg Glu Glu Leu Ser Pro Ala Ser Ser Ile Asn  
1 5 10 15  
Gly Cys Ser Thr Asp Gly Glu Ala Arg Gln Lys Lys Gly Pro Ala  
20 25 30  
40 Pro Arg Gln Gln Glu Glu Leu Cys Leu Val Cys Gly Asp Arg Ala Ser  
35 40 45  
Gly Tyr His Tyr Asn Ala Leu Thr Cys Glu Gly Cys Lys Gly Phe Phe  
50 55 60  
45 Arg Arg Ser Val Thr Lys Asn Ala Val Tyr Ile Cys Lys Phe Gly His  
65 70 75 80  
Ala Cys Glu Met Asp Met Tyr Met Arg Arg Lys Cys Gln Glu Cys Arg  
85 90 95  
50 Leu Lys Lys Cys Leu Ala Val Gly Met Arg Pro Glu Cys Val Val Pro  
100 105 110  
Glu Thr Gln Cys Ala Met Lys Arg Lys Glu Lys Lys Ala Gln Lys Glu  
115 120 125

EP 1 266 015 B1

Lys Asp Lys Leu Pro Val Ser Thr Thr Thr Val Asp Asp His Met Pro  
 130 135 140  
 Pro Ile Met Gln Cys Glu Pro Pro Pro Pro Glu Ala Ala Arg Ile His  
 145 150 160  
 Glu Val Val Pro Arg Phe Leu Ser Asp Lys Leu Leu Glu Thr Asn Arg  
 165 170 175  
 Gln Lys Asn Ile Pro Gln Leu Thr Ala Asn Gln Gln Phe Leu Ile Ala  
 180 185 190  
 Arg Leu Ile Trp Tyr Gln Asp Gly Tyr Glu Gln Pro Ser Asp Glu Asp  
 195 200 205  
 Leu Lys Arg Ile Thr Gln Thr Trp Gln Gln Ala Asp Asp Glu Asn Glu  
 210 215 220  
 Glu Ser Asp Thr Pro Phe Arg Gln Ile Thr Glu Met Thr Ile Leu Thr  
 225 230 240  
 Val Gln Leu Ile Val Glu Phe Ala Lys Gly Leu Pro Gly Phe Ala Lys  
 245 250 255  
 Ile Ser Gln Pro Asp Gln Ile Thr Leu Leu Lys Ala Cys Ser Ser Glu  
 260 265 270  
 Val Met Met Leu Arg Val Ala Arg Arg Tyr Asp Ala Ala Ser Asp Ser  
 275 280 285  
 Val Leu Phe Ala Asn Asn Gln Ala Tyr Thr Arg Asp Asn Tyr Arg Lys  
 290 295 300  
 Ala Gly Met Ala Tyr Val Ile Glu Asp Leu Leu His Phe Cys Arg Cys  
 305 310 320  
 Met Tyr Ser Met Ala Leu Asp Asn Ile His Tyr Ala Leu Leu Thr Ala  
 325 330 335  
 Val Val Ile Phe Ser Asp Arg Pro Gly Leu Glu Gln Pro Gln Leu Val  
 340 345 350  
 Glu Glu Ile Gln Arg Tyr Tyr Leu Asn Thr Leu Arg Ile Tyr Ile Leu  
 355 360 365  
 Asn Gln Leu Ser Gly Ser Ala Arg Ser Ser Val Ile Tyr Gly Lys Ile  
 370 375 380  
 Leu Ser Ile Leu Ser Glu Leu Arg Thr Leu Gly Met Gln Asn Ser Asn  
 385 390 400  
 Met Cys Ile Ser Leu Lys Leu Lys Asn Arg Lys Leu Pro Pro Phe Leu  
 405 410 415  
 Glu Glu Ile Trp Asp Val Ala Asp Met Ser His Thr Gln Pro Pro Pro  
 420 425 430  
 Ile Leu Glu Ser Pro Thr Asn Leu  
 435 440

<210> 64  
 <211> 943  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <221> misc feature  
 <223> Novel Sequence  
 <400> 64

atgacttcga aagtttatga tccagaacaa aggaaacgga tgataactgg tccgcagtgg 60  
 tgggccagat gtaaacaaat gaatgttctt gattcattta ttaattatta tgattcagaa 120  
 5 aaacatgcag aaaatgctgt tattttttta catggtaacg cggcctcttc ttattttatgg 180  
 cgacatgttg tgccacatat tgagccagta cgcgcgtgta ttataccaga ccttatttgg 240  
 atgggcaaat caggcaaatc tggtaatggt tcttataggt tacttgatca ttacaaatat 300  
 10 cttactgcat ggtttgaact tcttaattta ccaaagaaga tcatttttgt cggccatgat 360  
 tggggtgctt gtttggcatt tcattatagc tatgagcatc aagataagat caaagcaata 420  
 gttcacgctg aaagtgtagt agatgtgatt gaatcatggg atgaatggcc tgatattgaa 480  
 15 gaagatatgt cgttgatcaa atctgaagaa ggagaaaaaa tgggttttga gaataacttc 540  
 ttcgtggaac ccatgttgcc atcaaaaatc atgagaaagt tagaaccaga agaatttgca 600  
 gcatatcttg aaccattcaa agagaaaggt gaagttcgtc gtccaacatt atcatggcct 660  
 20 cgtgaaatcc cgttagtaaa aggtggttaa cctgacgttg tacaaattgt taggaattat 720  
 aatgcttacc tactgtcaag tgatgattha ccaaaaatgt ttattgaatc ggaccaggga 780  
 ttcctttcca atgctattgt tgaagggtgcc aagaagtttc ctaatactga atttgtaaaa 840  
 25 gtaaaaggtc ttcatTTTTc gcaagaagat gcacctgatg aaatgggaaa atatatcaaa 900  
 tcgttcgttg agcgagttct caaaaatgaa caataattct aga 943

## Claims

### 1. A gene expression modulation system comprising:

a) a first gene expression cassette that is capable of being expressed in a host cell comprising a polynucleotide sequence that encodes a first polypeptide comprising:

- i) a DNA-binding domain that recognizes a response element associated with a gene whose expression is to be modulated;
- ii) a ligand binding domain comprising a ligand binding domain from an ecdysone receptor polypeptide;

b) a second gene expression cassette that is capable of being expressed in the host cell comprising a polynucleotide sequence that encodes a second polypeptide comprising:

- i) a transactivation domain; and
- ii) a ligand binding domain comprising a ligand binding domain from a truncated retinoid X receptor comprising a deletion of at least the A and B domains;

wherein the transactivation domain is from a nuclear receptor other than an ecdysone receptor, a retinoid X receptor, or an ultraspiracle receptor, and wherein the ligand binding domains from the first polypeptide and the second polypeptide are different and dimerize.

### 2. The gene expression modulation system according to claim 1, further comprising a third gene expression cassette comprising:

- i) a response element to which the DNA-binding domain of the polypeptide binds;
- ii) a promoter that is activated by the transactivation domain of the second polypeptide; and
- iii) the gene whose expression is to be modulated.

3. The gene expression modulation system according to claim 1, wherein the ligand binding domain of the first polypeptide is encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, and SEQ ID NO: 10.
4. The gene expression modulation system according to claim 1, wherein the ligand binding domain of the first polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, and SEQ ID NO: 20.
5. The gene expression modulation system according to claim 1, wherein the ligand binding domain of the second polypeptide is encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, and SEQ ID NO: 30.
6. The gene expression modulation system according to claim 1, wherein the ligand binding domain of the second polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, and SEQ ID NO: 40.
7. A gene expression modulation system comprising:
  - a) a first gene expression cassette that is capable of being expressed in a host cell comprising a polynucleotide sequence that encodes a first polypeptide comprising:
    - i) a DNA-binding domain that recognizes a response element associated with a gene whose expression is to be modulated; and
    - ii) a ligand binding domain comprising a ligand binding domain from a truncated retinoid X receptor comprising a deletion of at least the A and B domains; and
  - b) a second gene expression cassette that is capable of being expressed in a host cell comprising a polynucleotide sequence that encodes a second polypeptide comprising:
    - i) a transactivation domain; and
    - ii) a ligand binding domain comprising a ligand binding domain from an ecdysone receptor;
 wherein the ligand binding domains from the first polypeptide and the second polypeptide are different and dimerize.
8. The gene expression modulation system according to claim 7, further comprising a third gene expression cassette comprising:
  - i) a response element to which the DNA-binding domain of the first polypeptide binds;
  - ii) a promoter that is activated by the transactivation domain of the second polypeptide; and
  - iii) the gene whose expression is being modulated.
9. The gene expression modulation system according to claim 7, wherein the ligand binding domain of the first polypeptide is encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, and SEQ ID NO: 30.
10. The gene expression modulation system according to claim 7, wherein the ligand binding domain of the first polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, and SEQ ID NO: 40.
11. The gene expression modulation system according to claim 7, wherein the ligand binding domain of the second polypeptide is encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ

ID NO: 8, SEQ ID NO: 9, and SEQ ID NO: 10.

12. The gene expression modulation system according to claim 7, wherein the ligand binding domain of the second polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, and SEQ ID NO: 20.

13. A gene expression cassette comprising a polynucleotide encoding a hybrid polypeptide selected from the group consisting of:

- a) a DNA-binding domain and an ecdysone receptor ligand binding domain, wherein the DNA binding domain is from a nuclear receptor other than an ecdysone receptor,
- b) a DNA-binding domain and a ligand binding domain from a truncated retinoid X receptor comprising a deletion of at least the A and B domains, wherein the DNA-binding domain is from a nuclear receptor other than a retinoid X receptor,
- c) a transactivation domain and an ecdysone receptor ligand binding domain, wherein the transactivation domain is from a nuclear receptor other than an ecdysone receptor, and
- d) a transactivation domain and a ligand binding domain from a truncated retinoid X receptor comprising a deletion of at least the A and B domains, wherein the transactivation domain is from a nuclear receptor other than a retinoid X receptor.

14. The gene expression cassette according to claim 13, wherein said expression cassette comprises a polynucleotide encoding a hybrid polypeptide comprising:

- a DNA-binding domain and an ecdysone receptor ligand binding domain wherein the DNA binding domain is from a nuclear receptor other than an ecdysone receptor; or
- a DNA-binding domain and a retinoid X receptor ligand binding domain, wherein the DNA-binding domain is from a nuclear receptor other than a retinoid X receptor; and

wherein the DNA-binding domain is GAL4 DNA-binding domain or a LexA DNA-binding domain.

15. The gene expression cassette according to claim 13, wherein said expression cassette comprises a polynucleotide encoding a hybrid polypeptide comprising:

- a transactivation domain and an ecdysone receptor ligand binding domain, wherein the transactivation domain is from a nuclear receptor other than an ecdysone receptor;
- or a transactivation domain and a retinoid X receptor ligand binding domain, wherein the transactivation domain is from a nuclear receptor other than a retinoid X receptor;

and wherein the transactivation domain is a VP16 transactivation domain.

16. A gene expression cassette comprising a polynucleotide encoding a hybrid polypeptide comprising a DNA-binding domain encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of GAL4 DBD (SEQ ID NO: 41) or a LexA DBD (SEQ ID NO: 43) and an ecdysone receptor ligand binding domain encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, and SEQ ID NO: 10.

17. A gene expression cassette comprising a polynucleotide encoding a hybrid polypeptide comprising a DNA-binding domain comprising an amino acid sequence selected from the group consisting of a GAL4 DBD (SEQ ID NO: 42) or a LexA DBD (SEQ ID NO: 44) and an ecdysone receptor ligand binding domain comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, and SEQ ID NO: 20.

18. A gene expression cassette comprising a polynucleotide encoding a hybrid polypeptide comprising a DNA-binding domain encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of a GAL4 DBD (SEQ ID NO: 41) or a LexA DBD (SEQ ID NO: 43) and a retinoid X receptor ligand binding domain encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO:

21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, and SEQ ID NO: 30.

19. A gene expression cassette comprising a polynucleotide encoding a hybrid polypeptide comprising a DNA-binding domain comprising an amino acid sequence selected from the group consisting of a GAL4 DBD (SEQ ID NO: 42) or a LexA DBD (SEQ ID NO: 44) and a retinoid X receptor ligand binding domain comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO 37, SEQ ID NO 38, SEQ ID NO: 39, and SEQ ID NO: 40.

20. A gene expression cassette comprising a polynucleotide encoding a hybrid polypeptide comprising a transactivation domain encoded by a polynucleotide comprising a nucleic acid sequence of SEQ ID NO: 45 and an ecdysone receptor ligand binding domain encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, and SEQ ID NO: 10.

21. A gene expression cassette comprising a polynucleotide encoding a hybrid polypeptide comprising a transactivation domain comprising an amino acid sequence of SEQ ID NO: 46 and an ecdysone receptor ligand binding domain comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, and SEQ ID NO: 20.

22. A gene expression cassette comprising a polynucleotide encoding a hybrid polypeptide comprising a transactivation domain encoded by a polynucleotide comprising a nucleic acid sequence of SEQ ID NO: 45 and a retinoid X receptor ligand binding domain encoded by a polynucleotide comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, and SEQ ID NO: 30.

23. A gene expression cassette comprising a polynucleotide encoding a hybrid polypeptide comprising a transactivation domain comprising an amino acid sequence of SEQ ID NO: 46 and a retinoid X receptor ligand binding domain comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO 37, SEQ ID NO 38, SEQ ID NO: 39, and SEQ ID NO: 40.

24. An isolated polynucleotide encoding a truncated retinoid X receptor polypeptide wherein the polynucleotide comprises a nucleic acid sequence selected from the group consisting of SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, and SEQ ID NO: 30.

25. An isolated truncated retinoid X receptor polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO 37, SEQ ID NO 38, SEQ ID NO: 39, and SEQ ID NO: 40.

26. An isolated polynucleotide comprising a polynucleotide sequence which encodes the isolated polypeptide according to claim 25.

27. A method of modulating the expression of a gene in a host cell comprising the gene to be modulated comprising the steps of:

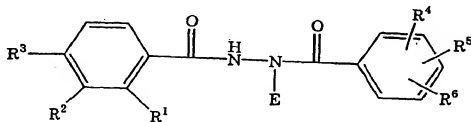
- a) introducing into the host cell the gene expression modulation system according to claim 1 or 2; and
- b) introducing into the host cell a ligand that independently combines with the ligand binding domains of the first polypeptide and the second polypeptide;

wherein the gene to be expressed is a component of a chimeric gene comprising:

- i) a response element to which the DNA binding domain from the first polypeptide binds;
- ii) a promoter that is activated by the transactivation domain of the second polypeptide; and
- iii) a gene whose expression is to be modulated.

whereby a complex is formed comprising the ligand, the first polypeptide, and the second polypeptide, and whereby the complex modulates expression of the gene in the host cell.

28. The method according to claim 27, wherein the ligand is a compound of the formula:



wherein:

E is a (C<sub>4</sub>-C<sub>6</sub>)alkyl containing a tertiary carbon or a cyano(C<sub>3</sub>-C<sub>6</sub>)alkyl containing a tertiary carbon:

R<sup>1</sup> is H, Me, Et, i-Pr, F, formyl, CF<sub>3</sub>, CHF<sub>2</sub>, CHCl<sub>2</sub>, CH<sub>2</sub>F, CH<sub>2</sub>Cl, CH<sub>2</sub>OH, CH<sub>2</sub>OMe, CH<sub>2</sub>CN, CN, <CCH>, 1-propynyl, 2-propynyl, vinyl, OH, OMe, OEt, cyclopropyl, CF<sub>2</sub>CF<sub>3</sub>, CH=CHCN, allyl, azido, SCN, or SCHF<sub>2</sub>; R<sub>2</sub> is H, Me, Et, n-Pr, i-Pr, formyl, CF<sub>3</sub>, CHF<sub>2</sub>, CHCl<sub>2</sub>, CH<sub>2</sub>F, CH<sub>2</sub>Cl, CH<sub>2</sub>OH, CH<sub>2</sub>OMe, CH<sub>2</sub>CN, CN, <CCH>, 1-propynyl, 2-propynyl, vinyl, Ac, F, Cl, OH, OMe, OEt, O-n-Pr, OAc, NMe<sub>2</sub>, NEt<sub>2</sub>, SMe, SET, SOCF<sub>3</sub>, OCF<sub>2</sub>CF<sub>2</sub>H, COEt, cyclopropyl, CF<sub>2</sub>CF<sub>3</sub>, CH=CHCN, allyl, azido, OCF<sub>3</sub>, OCHF<sub>2</sub>, O-i-Pr, SCN, SCHF<sub>2</sub>, SOME, NH-CN, or joined with R<sup>2</sup> and the phenyl carbons to which R<sup>2</sup> and R<sup>3</sup> are attached to form an ethylenedioxy, a dihydrofuryl ring with the oxygen adjacent to a phenyl carbon, or a dihydropyryl ring with the oxygen adjacent to a phenyl carbon; R<sup>3</sup> is H, Et, or joined with R<sup>2</sup> and the phenyl carbons to which R<sup>2</sup> and R<sup>3</sup> are attached to form an ethylenedioxy, a dihydrofuryl ring with the oxygen adjacent to a phenyl carbon, or a dihydropyryl ring with the oxygen adjacent to a phenyl carbon; R<sup>4</sup>, R<sup>5</sup>, and R<sup>6</sup> are independently H, Me, Et, F, Cl, Br, formyl, CF<sub>3</sub>, CHF<sub>2</sub>, CHCl<sub>2</sub>, CH<sub>2</sub>F, CH<sub>2</sub>Cl, CH<sub>2</sub>OH, CN, C<sup>o</sup>CH, 1-propynyl, 2-propynyl, vinyl, OMe, OEt, SMe, or SET..

29. An isolated host cell into which the gene expression modulation system according to claim 1 or 7 has been introduced.
30. The isolated host cell according to claim 29, wherein the host cell is selected from the group consisting of a bacterial cell, a fungal cell, a yeast cell, a plant cell, an animal cell, and a mammalian cell.
31. The isolated host cell according to claim 30, wherein the mammalian cell is a murine cell or a human cell.
32. A non-human organism comprising a host cell into which the gene expression modulation system according to claim 1 or 7 has been introduced.
33. The non-human organism according to claim 32, wherein the non-human organism is selected from the group consisting of a bacterium, a fungus, a yeast, a plant, an animal, and a mammal.
34. The non-human organism according to claim 33, wherein the mammal is selected from the group consisting of a mouse, a rat, a rabbit, a cat, a dog, a bovine, a goat, a pig, a horse, a sheep, a monkey, and a chimpanzee.

## Patentansprüche

1. Die Genexpression modulierendes System, welches folgendes umfaßt:

a) eine erste Genexpressionkassette, die in einer Wirtszelle exprimiert werden kann, umfassend eine Polynukleotidsequenz, die ein erstes Polypeptid codiert, welches folgendes umfaßt:



- i) eine DNA-Bindungsdomäne, die ein Antwoortelement, das mit einem Gen, dessen Expression moduliert werden soll, assoziiert ist, erkennt,
- ii) eine Ligandenbindungsdomäne, die eine Ligandenbindungsdomäne eines Ecdysonrezeptorpolypeptids umfaßt,

b) eine zweite Genexpressionskassette, die in der Wirtszelle exprimiert werden kann, umfassend eine Polynukleotidsequenz, die ein zweites Polypeptid codiert, welches folgendes umfaßt:

- i) eine Transaktivierungsdomäne und
- ii) eine Ligandenbindungsdomäne, die eine Ligandenbindungsdomäne eines verkürzten Retinoid X-Rezeptors, der eine Deletion wenigstens der A- und der B-Domäne aufweist, umfaßt,

wobei die Transaktivierungsdomäne von einem Kernrezeptor stammt, der kein Ecdysonrezeptor, kein Retinoid X-Rezeptor oder kein Ultraspiracle-Rezeptor ist, und wobei die Ligandenbindungsdomänen des ersten Polypeptids und des zweiten Polypeptids verschieden sind und dimerisieren.

2. Die Genexpression modulierendes System nach Anspruch 1, welches weiterhin eine dritte Genexpressionskassette umfaßt, die folgendes umfaßt:

- i) ein Antwoortelement, an das die DNA-Bindungsdomäne des Polypeptids bindet,
- ii) einen Promotor, der durch die Transaktivierungsdomäne des zweiten Polypeptids aktiviert wird, und
- iii) das Gen, dessen Expression moduliert werden soll.

3. Die Genexpression modulierendes System nach Anspruch 1, wobei die Ligandenbindungsdomäne des ersten Polypeptids durch ein Polynukleotid codiert wird, welches eine Nukleinsäuresequenz umfaßt, die aus der Gruppe ausgewählt ist, bestehend aus SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9 und SEQ ID NO:10.

4. Die Genexpression modulierendes System nach Anspruch 1, wobei die Ligandenbindungsdomäne des ersten Polypeptids eine Aminosäuresequenz umfaßt, die aus der Gruppe ausgewählt ist, bestehend aus SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19 und SEQ ID NO:20.

5. Die Genexpression modulierendes System nach Anspruch 1, wobei die Ligandenbindungsdomäne des zweiten Polypeptids durch ein Polynukleotid codiert wird, welches eine Nukleinsäuresequenz umfaßt, die aus der Gruppe ausgewählt ist, bestehend aus SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29 und SEQ ID NO:30.

6. Die Genexpression modulierendes System nach Anspruch 1, wobei die Ligandenbindungsdomäne des zweiten Polypeptids eine Aminosäuresequenz umfaßt, die aus der Gruppe ausgewählt ist, bestehend aus SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39 und SEQ ID NO:40.

7. Die Genexpression modulierendes System, welches folgendes umfaßt:

a) eine erste Genexpressionskassette, die in einer Wirtszelle exprimiert werden kann, umfassend eine Polynukleotidsequenz, welche ein erstes Polypeptid codiert, das folgendes umfaßt:

- i) eine DNA-Bindungsdomäne, die ein Antwoortelement, das mit einem Gen, dessen Expression moduliert werden soll, assoziiert ist, erkennt, und
- ii) eine Ligandenbindungsdomäne, die eine Ligandenbindungsdomäne eines verkürzten Retinoid X-Rezeptors, der eine Deletion wenigstens der A- und der B-Domäne aufweist, umfaßt und

b) eine zweite Genexpressionskassette, die in einer Wirtszelle exprimiert werden kann, umfassend eine Polynukleotidsequenz, die ein zweites Polypeptid codiert, das folgendes umfaßt:

- i) eine Transaktivierungsdomäne und
- ii) eine Ligandenbindungsdomäne, die eine Ligandenbindungsdomäne eines Ecdysonrezeptors umfaßt,

wobei die Ligandenbindungsdomänen des ersten Polypeptids und des zweiten Polypeptids verschieden sind und dimerisieren.

8. Die Genexpression modulierendes System nach Anspruch 7, welches weiterhin eine dritte Genexpressionskassette umfaßt, die folgendes umfaßt:

- i) ein Antwortelement, an das die DNA-Bindungsdomäne des ersten Polypeptids bindet,
- ii) einen Promotor, der durch die Transaktivierungsdomäne des zweiten Polypeptids aktiviert wird, und
- iii) das Gen, dessen Expression moduliert werden soll.

9. Die Genexpression modulierendes System nach Anspruch 7, wobei die Ligandenbindungsdomäne des ersten Polypeptids durch ein Polynukleotid codiert wird, welches eine Nukleinsäuresequenz umfaßt, die aus der Gruppe ausgewählt ist, bestehend aus SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29 und SEQ ID NO:30.

10. Die Genexpression modulierendes System nach Anspruch 7, wobei die Ligandenbindungsdomäne des ersten Polypeptids eine Aminosäuresequenz umfaßt, die aus der Gruppe ausgewählt ist, bestehend aus SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39 und SEQ ID NO:40.

11. Die Genexpression modulierendes System nach Anspruch 7, wobei die Ligandenbindungsdomäne des zweiten Polypeptids durch ein Polynukleotid codiert wird, welches eine Nukleinsäuresequenz umfaßt, die aus der Gruppe ausgewählt ist, bestehend aus SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9 und SEQ ID NO:10.

12. Die Genexpression modulierendes System nach Anspruch 7, wobei die Ligandenbindungsdomäne des zweiten Polypeptids eine Aminosäuresequenz umfaßt, die aus der Gruppe ausgewählt ist, bestehend aus SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19 und SEQ ID NO:20.

13. Genexpressionskassette, welche ein Polynukleotid umfaßt, das ein hybrides Polypeptid codiert, ausgewählt aus der Gruppe, bestehend aus:

- a) einer DNA-Bindungsdomäne und einer Ecdysonrezeptor-Ligandenbindungsdomäne, wobei die DNA-Bindungsdomäne von einem Kernrezeptor stammt, der kein Ecdysonrezeptor ist,
- b) einer DNA-Bindungsdomäne und einer Ligandenbindungsdomäne eines verkürzten Retinoid X-Rezeptors, welcher eine Deletion wenigstens der A- und der B-Domäne aufweist, wobei die DNA-Bindungsdomäne von einem Kernrezeptor stammt, der kein Retinoid X-Rezeptor ist,
- c) einer Transaktivierungsdomäne und einer Ecdysonrezeptor-Ligandenbindungsdomäne, wobei die Transaktivierungsdomäne von einem Kernrezeptor stammt, der kein Ecdysonrezeptor ist, und
- d) einer Transaktivierungsdomäne und einer Ligandenbindungsdomäne eines verkürzten Retinoid X-Rezeptors, welcher eine Deletion wenigstens der A- und der B-Domäne aufweist, wobei die Transaktivierungsdomäne von einem Kernrezeptor stammt, der kein Retinoid X-Rezeptor ist.

14. Genexpressionskassette nach Anspruch 13, wobei die Expressionskassette ein Polynukleotid umfaßt, welches ein hybrides Polypeptid codiert, das folgendes aufweist:

- eine DNA-Bindungsdomäne und eine Ecdysonrezeptor-Ligandenbindungsdomäne, wobei die DNA-Bindungsdomäne von einem Kernrezeptor stammt, der kein Ecdysonrezeptor ist,
- eine DNA-Bindungsdomäne und eine Retinoid X-Rezeptor-Ligandenbindungsdomäne, wobei die DNA-Bindungsdomäne von einem Kernrezeptor stammt, der kein Retinoid X-Rezeptor ist, und

wobei die DNA-Bindungsdomäne eine GAL4-DNA-Bindungsdomäne oder eine LexA-DNA-Bindungsdomäne ist.

15. Genexpressionskassette nach Anspruch 13, wobei die Expressionskassette ein Polynukleotid umfaßt, welches ein hybrides Polypeptid codiert, das folgendes aufweist:

- eine Transaktivierungsdomäne und eine Ecdysonrezeptor-Ligandenbindungsdomäne, wobei die Transaktivie-

rungsdomäne von einem Kernrezeptor stammt, der kein Ecdysonrezeptor ist, oder eine Transaktivierungsdomäne und eine Retinoid X-Rezeptor-Ligandenbindungsdomäne, wobei die Transaktivierungsdomäne von einem Kernrezeptor stammt, der kein Retinoid X-Rezeptor ist,

und wobei die Transaktivierungsdomäne eine VP16-Transaktivierungsdomäne ist.

16. Genexpressionskassette, welche ein Polynukleotid umfaßt, das ein hybrides Polypeptid codiert, welches eine DNA-Bindungsdomäne umfaßt, die durch ein Polynukleotid codiert ist, welches eine Nukleinsäuresequenz umfaßt, die aus der Gruppe ausgewählt ist, bestehend aus GAL4 DBD (SEQ ID NO:41) oder LexA DBD (SEQ ID NO:43), und eine Ecdysonrezeptor-Ligandenbindungsdomäne umfaßt, die durch ein Polynukleotid codiert ist, welches eine Nukleinsäuresequenz umfaßt, die aus der Gruppe ausgewählt ist, bestehend aus SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9 und SEQ ID NO:10.
17. Genexpressionskassette, welche ein Polynukleotid umfaßt, das ein hybrides Polypeptid codiert, welches eine DNA-Bindungsdomäne umfaßt, die eine Aminosäuresequenz umfaßt, ausgewählt aus der Gruppe, bestehend aus GAL4 DBD (SEQ ID NO:42) oder LexA DBD (SEQ ID NO:44), umfaßt, und eine Ecdysonrezeptor-Ligandenbindungsdomäne, die eine Aminosäuresequenz umfaßt, bestehend aus SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19 und SEQ ID NO:20.
18. Genexpressionskassette, welche ein Polynukleotid umfaßt, das ein hybrides Polypeptid codiert, welches eine DNA-Bindungsdomäne umfaßt, die durch ein Polynukleotid codiert wird, welches eine Nukleinsäuresequenz umfaßt, ausgewählt aus der Gruppe, bestehend aus GAL4 DBD (SEQ ID NO:41) oder LexA DBD (SEQ ID NO:43), und eine Retinoid X-Rezeptor-Ligandenbindungsdomäne umfaßt, die durch ein Polynukleotid codiert wird, welches eine Nukleinsäuresequenz umfaßt, ausgewählt aus der Gruppe, bestehend aus SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29 und SEQ ID NO:30.
19. Genexpressionskassette, welche ein Polynukleotid umfaßt, das ein hybrides Polypeptid codiert, welches eine DNA-Bindungsdomäne umfaßt, die eine Aminosäuresequenz umfaßt, ausgewählt aus der Gruppe, bestehend aus GAL4 DBD (SEQ ID NO:42) oder LexA DBD (SEQ ID NO:44), und eine Retinoid X-Rezeptor-Ligandenbindungsdomäne umfaßt, die eine Aminosäuresequenz umfaßt, ausgewählt aus der Gruppe, bestehend aus SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39 und SEQ ID NO:40.
20. Genexpressionskassette, welche ein Polynukleotid umfaßt, das ein hybrides Polypeptid codiert, welches eine Transaktivierungsdomäne umfaßt, die durch ein Polynukleotid codiert wird, welches eine Nukleinsäuresequenz der SEQ ID NO:45 umfaßt, und eine Ecdysonrezeptor-Ligandenbindungsdomäne umfaßt, die durch ein Polynukleotid codiert wird, welches eine Nukleinsäuresequenz umfaßt, ausgewählt aus der Gruppe, bestehend aus SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9 und SEQ ID NO:10.
21. Genexpressionskassette, welche ein Polynukleotid umfaßt, das ein hybrides Polypeptid codiert, welches eine Transaktivierungsdomäne umfaßt, die eine Aminosäuresequenz der SEQ ID NO:46 umfaßt, und eine Ecdysonrezeptor-Ligandenbindungsdomäne umfaßt, die eine Aminosäuresequenz umfaßt, ausgewählt aus der Gruppe, bestehend aus SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19 und SEQ ID NO:20.
22. Genexpressionskassette, welche ein Polynukleotid umfaßt, das ein hybrides Polypeptid codiert, welches eine Transaktivierungsdomäne umfaßt, die durch ein Polynukleotid codiert wird, welches eine Nukleinsäuresequenz der SEQ ID NO:45 umfaßt, und eine Retinoid X-Rezeptor-Ligandenbindungsdomäne umfaßt, die durch ein Polynukleotid codiert wird, welches eine Nukleinsäuresequenz umfaßt, ausgewählt aus der Gruppe, bestehend aus SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29 und SEQ ID NO:30.
23. Genexpressionskassette, welche ein Polynukleotid umfaßt, das ein hybrides Polypeptid codiert, welches eine Transaktivierungsdomäne umfaßt, die eine Aminosäuresequenz der SEQ ID NO:46 umfaßt, und eine Retinoid X-Rezeptor-Ligandenbindungsdomäne umfaßt, die eine Aminosäuresequenz umfaßt, ausgewählt aus der Gruppe, bestehend

aus SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39 und SEQ ID NO:40.

24. Isoliertes Polypeptid, welches ein verkürztes Retinoid X-Rezeptorpolypeptid codiert, wobei das Polynukleotid eine Nukleinsäuresequenz umfaßt, ausgewählt aus der Gruppe, bestehend aus SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29 und SEQ ID NO:30.

25. Isoliertes verkürztes Retinoid X-Rezeptorpolypeptid, welches eine Aminosäuresequenz umfaßt, ausgewählt aus der Gruppe, bestehend aus SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39 und SEQ ID NO:40.

26. Isoliertes Polynukleotid, welches eine Polynukleotidsequenz umfaßt, die das isolierte Polypeptid nach Anspruch 25 codiert.

27. Verfahren zum Modulieren der Expression eines Gens in einer Wirtszelle, die das zu modulierende Gen enthält, welches die folgenden Stufen umfaßt:

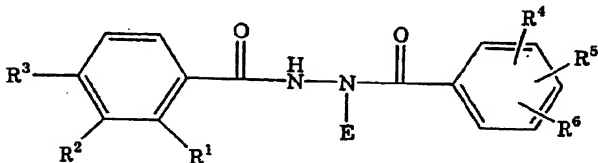
- Einbringen des die Genexpression modulierenden Systems nach einem der Ansprüche 1 oder 2 in die Wirtszelle und
- Einbringen eines Liganden, der sich unabhängig mit den Ligandenbindungsdomänen des ersten Polypeptids und des zweiten Polypeptids vereinigt, in die Wirtszelle,

wobei das zu exprimierende Gen eine Komponente eines chimären Gens ist, welches folgendes umfaßt:

- ein Antwortelement, an das die DNA-Bindungsdomäne des ersten Polypeptids bindet,
- einen Promotor, der durch die Transaktivierungsdomäne des zweiten Polypeptids aktiviert wird, und
- das Gen, dessen Expression moduliert werden soll,

wobei ein Komplex gebildet wird, der den Liganden, das erste Polypeptid und das zweite Polypeptid umfaßt, und wobei der Komplex die Expression des Gens in der Wirtszelle moduliert.

28. Verfahren nach Anspruch 27, wobei der Ligand eine Verbindung der Formel



ist, wobei:

E ein  $(C_4-C_6)$ -Alkyl, das ein tertiäres Kohlenstoffatom enthält, oder ein Cyano- $(C_3-C_5)$ -alkyl, das ein tertiäres Kohlenstoffatom enthält, ist,  
 $R^1$  H, Me, Et, i-Pr, F, Formyl,  $CF_3$ ,  $CHF_2$ ,  $CHCl_2$ ,  $CH_2F$ ,  $CH_2Cl$ ,  $CH_2OH$ ,  $CH_2OMe$ ,  $CH_2CN$ , CN, CCH, 1-Propinyl, 2-Propinyl, Vinyl, OH, OMe, OEt, Cyclopropyl,  $CF_2CF_3$ ,  $CH=CHCN$ , Allyl, Azido, SCN oder  $SCHF_2$  ist,  
 $R^2$  H, Me, Et, n-Pr, i-Pr, Formyl,  $CF_3$ ,  $CHF_2$ ,  $CHCl_2$ ,  $CH_2F$ ,  $CH_2Cl$ ,  $CH_2OH$ ,  $CH_2OMe$ ,  $CH_2CN$ , CN, CCH, 1-Propinyl, 2-Propinyl, Vinyl, Ac, F, Cl, OH, OMe, OEt, O-n-Pr, OAc,  $NMe_2$ ,  $NEt_2$ , SMe, SEt,  $SOCF_3$ ,  $OCF_2CF_2H$ , COEt, Cyclopropyl,  $CF_2CF_3$ ,  $CH=CHCN$ , Allyl, Azido,  $OCF_3$ ,  $OCHF_2$ , 0-i-Pr, SCN,  $SCHF_2$ , SMe, NH-CN ist oder mit  $R^3$  und den Phenylkohlenstoffen, an die  $R^2$  und  $R^3$  angebunden sind, unter Bildung von Ethylendioxy, eines Dihydrofurylrings, bei dem der Sauerstoff zu einem Phenylkohlenstoff benachbart ist, oder eines Dihy-

dropyrylring, bei dem der Sauerstoff zu einem Phenylkohlenstoff benachbart ist, verbunden ist,  $R^3$  H, Et oder mit  $R^2$  und den Phenylkohlenstoffen, an die  $R^2$  und  $R^3$  angebunden sind, unter Bildung von Ethylendioxy, eines Dihydrofurylring, bei dem der Sauerstoff zu einem Phenylkohlenstoff benachbart ist, oder eines Dihydrodropyrylring, bei dem der Sauerstoff zu einem Phenylkohlenstoff benachbart ist, verbunden ist,  $R^4$ ,  $R^5$  und  $R^6$  unabhängig voneinander H, Me, Et, F, Cl, Br, Formyl,  $CF_3$ ,  $CHF_2$ ,  $CHCl_2$ ,  $CH_2F$ ,  $CH_2Cl$ ,  $CH_2OH$ , CN,  $C^*CH$ , 1-Propinyl, 2-Propinyl, Vinyl, OMe, OEt, SMe oder SEt sind.

29. Isolierte Wirtszelle, in die das die Genexpression modulierende System nach Anspruch 1 oder 7 eingebracht wurde.

30. Isolierte Wirtszelle nach Anspruch 29, wobei die Wirtszelle aus der Gruppe ausgewählt ist, bestehend aus einer Bakterienzelle, einer Pilzzelle, einer Hefezelle, einer Pflanzenzelle, einer Tierzelle und einer Säugerzelle.

31. Isolierte Wirtszelle nach Anspruch 30, wobei die Säugerzelle eine Mauszelle oder eine menschliche Zelle ist.

32. Nicht-menschlicher Organismus, welcher eine Wirtszelle, in die das die Genexpression modulierende System nach Anspruch 1 oder 7 eingebracht wurde, umfaßt.

33. Nicht-menschlicher Organismus nach Anspruch 32, wobei der nicht-menschliche Organismus aus der Gruppe ausgewählt ist, bestehend aus einem Bakterium, einem Pilz, einer Hefe, einer Pflanze, einem Tier und einem Säuger.

34. Nicht-menschlicher Organismus nach Anspruch 33, wobei der Säuger aus der Gruppe ausgewählt ist, bestehend aus einer Maus, einer Ratte, einem Kaninchen, einer Katze, einem Hund, einem Rind, einer Ziege, einem Schwein, einem Pferd, einem Schaf, einem Affen und einem Schimpansen.

## Revendications

1. Système de modulation d'expression génique comprenant :

a) une première cassette d'expression génique qui est capable d'être exprimée dans une cellule hôte comprenant une séquence polynucléotidique qui code pour un premier polypeptide comprenant :

i) un domaine de liaison à l'ADN qui reconnaît un élément de réponse associé à un gène dont l'expression doit être modulée ;

ii) un domaine de liaison au ligand comprenant un domaine de liaison au ligand issu d'un polypeptide du récepteur de l'ecdysone ;

b) une deuxième cassette d'expression génique qui est capable d'être exprimée dans la cellule hôte comprenant une séquence polynucléotidique qui code pour un second polypeptide comprenant :

i) un domaine de transactivation ; et

ii) un domaine de liaison au ligand comprenant un domaine de liaison au ligand issu d'un récepteur X des rétinoïdes tronqué comprenant une délétion d'au moins les domaines A et B ;

dans lequel le domaine de transactivation provient d'un récepteur nucléaire autre qu'un récepteur de l'ecdysone, d'un récepteur X des rétinoïdes ou d'un récepteur ultraspiracle ; et dans lequel les domaines de liaison au ligand issus du premier polypeptide et du second polypeptide sont différents et se dimérisent.

2. Système de modulation d'expression génique selon la revendication 1, comprenant de plus une troisième cassette d'expression génique comprenant :

i) un élément de réponse auquel se lie le domaine de liaison à l'ADN du polypeptide ;

ii) un promoteur qui est activé par le domaine de transactivation du second polypeptide ; et

iii) le gène dont l'expression doit être modulée.

3. Système d'expression génique selon la revendication 1, dans lequel le domaine de liaison au ligand du premier polypeptide est codé par un polynucléotide comprenant une séquence d'acide nucléique choisie dans le groupe formé par SEQ ID N° 1, SEQ ID N° 2, SEQ ID N° 3, SEQ ID N° 4, SEQ ID N° 5, SEQ ID N° 6, SEQ ID N° 7, SEQ

ID N° 8, SEQ ID N° 9 et SEQ ID N° 10.

4. Système de modulation d'expression génique selon la revendication 1, dans lequel le domaine de liaison au ligand du premier polypeptide comprend une séquence d'acides aminés choisie dans le groupe formé par SEQ ID N° 11, SEQ ID N° 12, SEQ ID N° 13, SEQ ID N° 14, SEQ ID N° 15, SEQ ID N° 16, SEQ ID N° 17, SEQ ID N° 18, SEQ ID N° 19 et SEQ ID N° 20.
5. Système de modulation d'expression génique selon la revendication 1, dans lequel le domaine de liaison au ligand du second polypeptide est codé par un polynucléotide comprenant une séquence d'acide nucléique choisie dans le groupe formé par SEQ ID N° 21, SEQ ID N° 22, SEQ ID N° 23, SEQ ID N° 24, SEQ ID N° 25, SEQ ID N° 26, SEQ ID N° 27, SEQ ID N° 28, SEQ ID N° 29 et SEQ ID N° 30.
6. Système de modulation d'expression génique selon la revendication 1, dans lequel le domaine de liaison au ligand du second polypeptide comprend une séquence d'acides aminés choisie dans le groupe formé par SEQ ID N° 31, SEQ ID N° 32, SEQ ID N° 33, SEQ ID N° 34, SEQ ID N° 35, SEQ ID N° 36, SEQ ID N° 37, SEQ ID N° 38, SEQ ID N° 39 et SEQ ID N° 40.
7. Système de modulation d'expression génique comprenant :
  - a) une première cassette d'expression génique qui est capable d'être exprimée dans une cellule hôte comprenant une séquence polynucléotidique qui code pour un premier polypeptide comprenant :
    - i) un domaine de liaison à l'ADN qui reconnaît un élément de réponse associé à un gène dont l'expression doit être modulée ; et
    - ii) un domaine de liaison au ligand comprenant un domaine de liaison au ligand issu d'un récepteur X des rétinoïdes tronqué comprenant une délétion d'au moins les domaines A et B ; et
  - b) une deuxième cassette d'expression génique qui est capable d'être exprimée dans une cellule hôte comprenant une séquence polynucléotidique qui code pour un second polypeptide comprenant :
    - i) un domaine de transactivation ; et
    - ii) un domaine de liaison au ligand comprenant un domaine de liaison au ligand issu d'un récepteur de l'ecdysone ;
8. Système de modulation d'expression génique selon la revendication 7, comprenant de plus une troisième cassette d'expression génique comprenant :
  - i) un élément de réponse auquel se lie le domaine de liaison à l'ADN du premier polypeptide ;
  - ii) un promoteur qui est activé par le domaine de transactivation du second polypeptide ; et
  - iii) le gène dont l'expression est modulée.
9. Système de modulation d'expression génique selon la revendication 7, dans lequel le domaine de liaison au ligand du premier polypeptide est codé par un polynucléotide comprenant une séquence d'acide nucléique choisie dans le groupe formé par SEQ ID N° 21, SEQ ID N° 22, SEQ ID N° 23, SEQ ID N° 24, SEQ ID N° 25, SEQ ID N° 26, SEQ ID N° 27, SEQ ID N° 28, SEQ ID N° 29 et SEQ ID N° 30.
10. Système de modulation d'expression génique selon la revendication 7, dans lequel le domaine de liaison au ligand du premier polypeptide comprend une séquence d'acides aminés choisie dans le groupe formé par SEQ ID N° 31, SEQ ID N° 32, SEQ ID N° 33, SEQ ID N° 34, SEQ ID N° 35, SEQ ID N° 36, SEQ ID N° 37, SEQ ID N° 38, SEQ ID N° 39 et SEQ ID N° 40.
11. Système de modulation d'expression génique selon la revendication 7, dans lequel le domaine de liaison au ligand du second polypeptide est codé par un polynucléotide comprenant une séquence d'acide nucléique choisie dans le groupe formé par SEQ ID N° 1, SEQ ID N° 2, SEQ ID N° 3, SEQ ID N° 4, SEQ ID N° 5, SEQ ID N° 6, SEQ ID N° 7, SEQ ID N° 8, SEQ ID N° 9 et SEQ ID N° 10.

12. Système de modulation d'expression génique selon la revendication 7, dans lequel le domaine de liaison au ligand du second polypeptide comprend une séquence d'acides aminés choisie dans le groupe formé par SEQ ID N° 11, SEQ ID N° 12, SEQ ID N° 13, SEQ ID N° 14, SEQ ID N° 15, SEQ ID N° 16, SEQ ID N° 17, SEQ ID N° 18, SEQ ID N° 19 et SEQ ID N° 20.

13. Cassette d'expression génique comprenant un polynucléotide codant pour un polypeptide hybride choisi dans le groupe formé par :

- a) un domaine de liaison à l'ADN et un domaine de liaison au ligand de récepteur de l'ecdysone, le domaine de liaison à l'ADN étant issu d'un récepteur nucléaire autre qu'un récepteur de l'ecdysone ;
- b) un domaine de liaison à l'ADN et un domaine de liaison au ligand issu d'un récepteur X des rétinoïdes tronqué comprenant une délétion d'au moins les domaines A et B, le domaine de liaison à l'ADN étant issu d'un récepteur nucléaire autre qu'un récepteur X des rétinoïdes ;
- c) un domaine de transactivation et un domaine de liaison au ligand de récepteur de l'ecdysone, le domaine de transactivation étant issu d'un récepteur nucléaire autre qu'un récepteur de l'ecdysone ; et
- d) un domaine de transactivation et un domaine de liaison au ligand issu d'un récepteur X des rétinoïdes tronqué comprenant une délétion d'au moins les domaines A et B, le domaine de transactivation étant issu d'un récepteur nucléaire autre qu'un récepteur X des rétinoïdes.

14. Cassette d'expression génique selon la revendication 13, dans laquelle ladite cassette d'expression génique comprend un polynucléotide codant pour un polypeptide hybride comprenant :

- un domaine de liaison à l'ADN et un domaine de liaison au ligand de récepteur de l'ecdysone, le domaine de liaison à l'ADN étant issu d'un récepteur nucléaire autre qu'un récepteur de l'ecdysone ; ou
- un domaine de liaison à l'ADN et un domaine de liaison au ligand de récepteur X des rétinoïdes, le domaine de liaison à l'ADN étant issu d'un récepteur nucléaire autre qu'un récepteur X des rétinoïdes ; et

dans laquelle le domaine de liaison à l'ADN est un domaine de liaison à l'ADN de GAL4 ou un domaine de liaison à l'ADN de LexA.

15. Cassette d'expression génique selon la revendication 13, dans laquelle ladite cassette d'expression génique comprend un polynucléotide codant pour un polypeptide hybride comprenant :

- un domaine de transactivation et un domaine de liaison au ligand de récepteur de l'ecdysone, le domaine de transactivation étant issu d'un récepteur nucléaire autre qu'un récepteur de l'ecdysone ;
- ou un domaine de transactivation et un domaine de liaison au ligand de récepteur X des rétinoïdes, le domaine de transactivation étant issu d'un récepteur nucléaire autre qu'un récepteur X des rétinoïdes ;

et dans laquelle le domaine de transactivation est un domaine de transactivation de VP16.

16. Cassette d'expression génique comprenant un polynucléotide codant pour un polypeptide hybride comprenant un domaine de liaison à l'ADN codé par un polynucléotide comprenant une séquence d'acide nucléique choisie dans le groupe formé par un DBD de GAL4 (SEQ ID N° 41) ou un DBD de LexA (SEQ ID N° 43) et un domaine de liaison au ligand de récepteur de l'ecdysone codé par un polynucléotide comprenant une séquence d'acide nucléique choisie dans le groupe formé par SEQ ID N° 1, SEQ ID N° 2, SEQ ID N° 3, SEQ ID N° 4, SEQ ID N° 5, SEQ ID N° 6, SEQ ID N° 7, SEQ ID N° 8, SEQ ID N° 9 et SEQ ID N° 10.

17. Cassette d'expression génique comprenant un polynucléotide codant pour un polypeptide hybride comprenant un domaine de liaison à l'ADN comprenant une séquence d'acides aminés choisie dans le groupe formé par un DBD de GAL4 (SEQ ID N° 42) ou un DBD de LexA (SEQ ID N° 44) et un domaine de liaison au ligand de récepteur de l'ecdysone comprenant une séquence d'acides aminés choisie dans le groupe formé par SEQ ID N° 11, SEQ ID N° 12, SEQ ID N° 13, SEQ ID N° 14, SEQ ID N° 15, SEQ ID N° 16, SEQ ID N° 17, SEQ ID N° 18, SEQ ID N° 19 et SEQ ID N° 20.

18. Cassette d'expression génique comprenant un polynucléotide codant pour un polypeptide hybride comprenant un domaine de liaison à l'ADN codé par un polynucléotide comprenant une séquence d'acide nucléique choisie dans le groupe formé par un DBD de GAL4 (SEQ ID N° 41) ou un DBD de LexA (SEQ ID N° 43) et un domaine de liaison au ligand de récepteur X des rétinoïdes codé par un polynucléotide comprenant une séquence d'acide nucléique

choisie dans le groupe formé par SEQ ID N° 21, SEQ ID N° 22, SEQ ID N° 23, SEQ ID N° 24, SEQ ID N° 25, SEQ ID N° 26, SEQ ID N° 27, SEQ ID N° 28, SEQ ID N° 29 et SEQ ID N° 30.

19. Cassette d'expression génique comprenant un polynucléotide codant pour un polypeptide hybride comprenant un domaine de liaison à l'ADN comprenant une séquence d'acides aminés choisie dans le groupe formé par un DBD de GAL4 (SEQ ID N° 42) ou un DBD de LexA (SEQ ID N° 44) et un domaine de liaison au ligand de récepteur X des rétinoïdes comprenant une séquence d'acides aminés choisie dans le groupe formé par SEQ ID N° 31, SEQ ID N° 32, SEQ ID N° 33, SEQ ID N° 34, SEQ ID N° 35, SEQ ID N° 36, SEQ ID N° 37, SEQ ID N° 38, SEQ ID N° 39 et SEQ ID N° 40.

20. Cassette d'expression génique comprenant un polynucléotide codant pour un polypeptide hybride comprenant un domaine de transactivation codé par un polynucléotide comprenant une séquence d'acide nucléique de SEQ ID N° 45 et un domaine de liaison au ligand de récepteur de l'ecdysone codé par un polynucléotide comprenant une séquence d'acide nucléique choisie dans le groupe formé par SEQ ID N° 1, SEQ ID N° 2, SEQ ID N° 3, SEQ ID N° 4, SEQ ID N° 5, SEQ ID N° 6, SEQ ID N° 8, SEQ ID N° 9 et SEQ ID N° 10.

21. Cassette d'expression génique comprenant un polynucléotide codant pour un polypeptide hybride comprenant un domaine de transactivation comprenant une séquence d'acides aminés de SEQ ID N° 46 et un domaine de liaison au ligand de récepteur de l'ecdysone comprenant une séquence d'acides aminés choisie dans le groupe formé par SEQ ID N° 11, SEQ ID N° 12, SEQ ID N° 13, SEQ ID N° 14, SEQ ID N° 15, SEQ ID N° 16, SEQ ID N° 17, SEQ ID N° 18, SEQ ID N° 19 et SEQ ID N° 20.

22. Cassette d'expression génique comprenant un polynucléotide codant pour un polypeptide hybride comprenant un domaine de transactivation codé par un polynucléotide comprenant une séquence d'acide nucléique de SEQ ID N° 45 et un domaine de liaison au ligand de récepteur X des rétinoïdes codé par un polynucléotide comprenant une séquence d'acide nucléique choisie dans le groupe formé par SEQ ID N° 21, SEQ ID N° 22, SEQ ID N° 23, SEQ ID N° 24, SEQ ID N° 25, SEQ ID N° 26, SEQ ID N° 27, SEQ ID N° 28, SEQ ID N° 29 et SEQ ID N° 30.

23. Cassette d'expression génique comprenant un polynucléotide codant pour un polypeptide hybride comprenant un domaine de transactivation comprenant une séquence d'acides aminés de SEQ ID N° 46 et un domaine de liaison au ligand de récepteur X des rétinoïdes comprenant une séquence d'acides aminés choisie dans le groupe formé par SEQ ID N° 31, SEQ ID N° 32, SEQ ID N° 33, SEQ ID N° 34, SEQ ID N° 35, SEQ ID N° 36, SEQ ID N° 37, SEQ ID N° 38, SEQ ID N° 39 et SEQ ID N° 40.

24. Polynucléotide isolé codant pour un polypeptide de récepteur X des rétinoïdes tronqué, dans lequel le polynucléotide comprend une séquence d'acide nucléique choisie dans le groupe formé par SEQ ID N° 21, SEQ ID N° 22, SEQ ID N° 23, SEQ ID N° 24, SEQ ID N° 25, SEQ ID N° 26, SEQ ID N° 27, SEQ ID N° 28, SEQ ID N° 29 et SEQ ID N° 30.

25. Polypeptide de récepteur X des rétinoïdes tronqué isolé comprenant une séquence d'acides aminés choisie dans le groupe formé par SEQ ID N° 31, SEQ ID N° 32, SEQ ID N° 33, SEQ ID N° 34, SEQ ID N° 35, SEQ ID N° 36, SEQ ID N° 37, SEQ ID N° 38, SEQ ID N° 39 et SEQ ID N° 40.

26. Polynucléotide isolé comprenant une séquence polynucléotidique qui code pour le polypeptide isolé selon la revendication 25.

27. Procédé de modulation de l'expression d'un gène dans une cellule hôte comprenant le gène à moduler, comprenant les étapes de :

- a) introduction dans la cellule hôte du système de modulation d'expression génique selon la revendication 1 ou 2 ; et
- b) introduction dans la cellule hôte d'un ligand qui se combine indépendamment avec les domaines de liaison au ligand du premier polypeptide et du second polypeptide ;

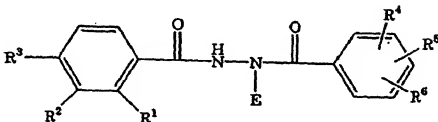
dans lequel le gène à exprimer est un composant d'un gène chimère comprenant :

- i) un élément de réponse auquel se lie le domaine de liaison à l'ADN issu du premier polypeptide ;
- ii) un promoteur qui est activé par le domaine de transactivation du second polypeptide ; et
- iii) un gène dont l'expression doit être modulée,



en sorte que soit formé un complexe comprenant le ligand, le premier polypeptide et le second polypeptide, et en sorte que le complexe module l'expression du gène dans la cellule hôte.

28. Procédé selon la revendication 27, dans lequel le ligand est un composé de formule :



dans laquelle :

E est un groupe alkyle en C<sub>4</sub> à C<sub>6</sub> contenant un atome de carbone tertiaire ou un groupe cyano(alkyle en C<sub>3</sub> à C<sub>5</sub>) contenant un atome de carbone tertiaire ;

R<sup>1</sup> est H, Me, Et, i-Pr, F, un groupe formyle, CF<sub>3</sub>, CHF<sub>2</sub>, CHCl<sub>2</sub>, CH<sub>2</sub>F, CH<sub>2</sub>Cl, CH<sub>2</sub>OH, CH<sub>2</sub>OMe, CH<sub>2</sub>CN, CN, CCH, 1-propynyle, 2-propynyle, vinyle, OH, OMe, OEt, cyclopropyle, CF<sub>2</sub>CF<sub>3</sub>, CH=CHCN, allyle, azido, SCN ou SCHF<sub>2</sub> ;

R<sup>2</sup> est H, Me, Et, n-Pr, i-Pr, un groupe formyle, CF<sub>3</sub>, CHF<sub>2</sub>, CHCl<sub>2</sub>, CH<sub>2</sub>F, CH<sub>2</sub>Cl, CH<sub>2</sub>OH, CH<sub>2</sub>OMe, CH<sub>2</sub>CN, CN, CCH, 1-propynyle, 2-propynyle, vinyle, Ac, F, Cl, OH, OMe, OEt, O-n-Pr, OAc, NMe<sub>2</sub>, NEt<sub>2</sub>, SMe, SET, SOCF<sub>3</sub>, OCF<sub>2</sub>CF<sub>2</sub>H, COEt, cyclopropyle, CF<sub>2</sub>CF<sub>3</sub>, CH=CHCN, allyle, azido, OCF<sub>3</sub>, OCHF<sub>2</sub>, O-i-Pr, SCN, SCHF<sub>2</sub>, SOMe, NH-CN, ou est joint à R<sup>3</sup> et aux atomes de carbone phéniliques auxquels R<sup>2</sup> et R<sup>3</sup> sont attachés pour former un groupe éthylènedioxy, un cycle dihydrofuryle dont l'atome d'oxygène est adjacent à un atome de carbone phénilique, ou un cycle dihydropyryle dont l'atome d'oxygène est adjacent à un atome de carbone phénilique ;

R<sup>3</sup> est H, Et, ou est joint à R<sup>2</sup> et aux atomes de carbone phéniliques auxquels R<sup>2</sup> et R<sup>3</sup> sont attachés pour former un groupe éthylènedioxy, un cycle dihydrofuryle dont l'oxygène est adjacent à un atome de carbone phénilique, ou un cycle dihydropyryle dont l'oxygène est adjacent à un atome de carbone phénilique ;

R<sup>4</sup>, R<sup>5</sup> et R<sup>6</sup> sont indépendamment H, Me, Et, F, Cl, Br, un groupe formyle, CF<sub>3</sub>, CHF<sub>2</sub>, CHCl<sub>2</sub>, CH<sub>2</sub>F, CH<sub>2</sub>Cl, CH<sub>2</sub>OH, CN, C<sup>o</sup>CH, 1-propynyle, 2-propynyle, vinyle, OMe, OEt, SMe ou SET.

29. Cellule hôte isolée dans laquelle a été introduit le système de modulation d'expression génique selon la revendication 1 ou 7.

30. Cellule hôte isolée selon la revendication 29, dans laquelle la cellule hôte est choisie dans le groupe formé par une cellule bactérienne, une cellule fongique, une cellule de levure, une cellule végétale, une cellule animale et une cellule de mammifère.

31. Cellule hôte isolée selon la revendication 30, dans laquelle la cellule de mammifère est une cellule murine ou une cellule humaine.

32. Organisme non humain comprenant une cellule hôte dans laquelle a été introduit le système de modulation d'expression génique selon la revendication 1 ou 7.

33. Organisme non humain selon la revendication 32, dans lequel l'organisme non humain est choisi dans le groupe formé par une bactérie, un champignon, une levure, une plante, un animal et un mammifère.

34. Organisme non humain selon la revendication 33, dans lequel le mammifère est choisi dans le groupe formé par une souris, un rat, un lapin, un chat, un chien, un bovin, une chèvre, un porc, un cheval, un mouton, un singe et un chimpanzé.

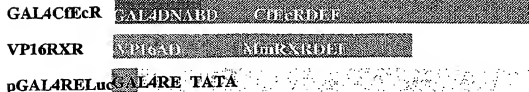


Figure 1

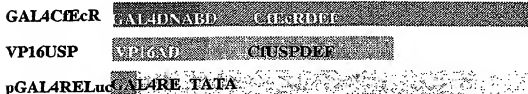


Figure 2

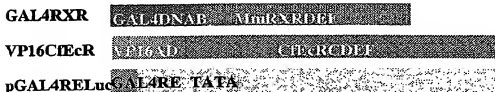


Figure 3

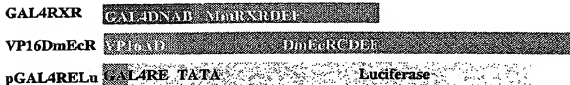


Figure 4

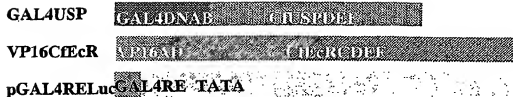


Figure 5

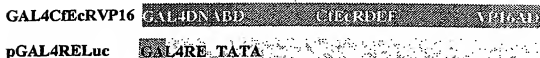


Figure 6



Figure 7

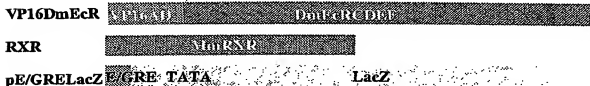


Figure 8



Figure 9



Figure 10

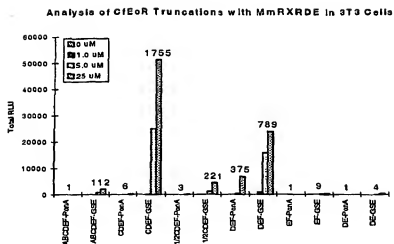


Figure 11

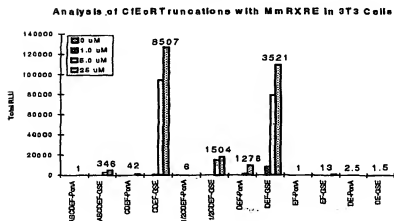


Figure 12

## Analysis of MmRXR Truncations with C1EcRCDEF in 3T3 Cells

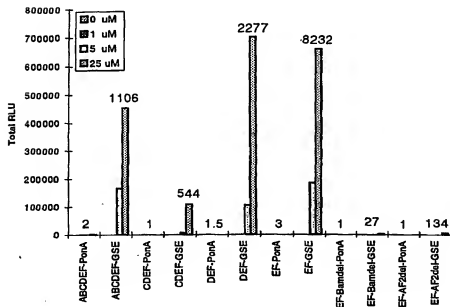


Figure 13

## Analysis of MmRXR Truncations with C1EcRDEF in 3T3 Cells

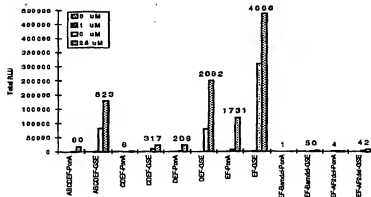


Figure 14

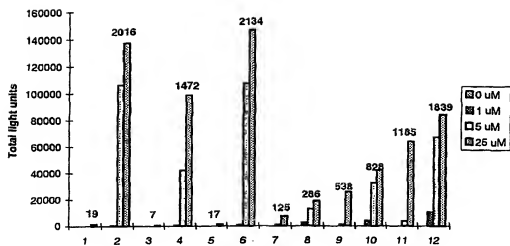


Figure 15